The

Geophysical Journal

of the

Royal Astronomical Society

incorporating

the Geophysical Supplement to the Monthly Notices of the R. A. S.

Volume 4, 1961

A special volume of invited contributions dedicated to

Sir Harold Jeffreys

by some of his students and colleagues on the occasion of his 70th birthday

Published and Sold by the

Royal Astronomical Society

Burlington House, London, W.1, England

1961

The Geophysical Journal

of the Royal Astronomical Society

Editor

S. K. Runcorn, M.A., Ph.D., F.R.A.S.

Published for the Royal Astronomical Society by Pergamon Press Limited, London, New York, Paris, Frankfurt

The Geophysical Journal is devoted to the publication of research in geophysics and related subjects. It includes original papers, short notes and letters to the editor as well as invited review papers.

The aims of the Journal are to publish papers of a high standard, to give them a wide international circulation and to ensure rapidity of publication.

Each year several reports of geophysical meetings are also published and the Editors are always glad to hear of such meetings and to receive short reports of them.

Notice to Authors

Authors of papers offered for publication need not be Fellows of the Society.

No charges of any kind are made for printing papers or articles accepted. Contributors are supplied on request with 25 reprints free and greater numbers are obtainable at moderate rates, e.g. 100 copies of a 12-page reprint (exc. plates) cost £11 14s. (U.S. $32). Special arrangements can be made for institutions requiring reprints of their contributions with covers, special pagination etc., for issue in their own series of publications.

Attention is drawn to the recommendations to authors printed on the inside of the back cover.

Volume 5 No. 1 May 1961

The following contributions have been accepted for publication in the above issue:

T. Rikitake, The effect of the ocean on rapid geomagnetic changes.

K. M. Creer, Superparamagnetism in red sandstones.

A. H. Cook, The comparison of the Earth's gravitational potential derived from satellite observations with gravity observations on the surface.

R. W. Girdler, The measurement and computation of anisotropy of magnetic susceptibility of rocks.

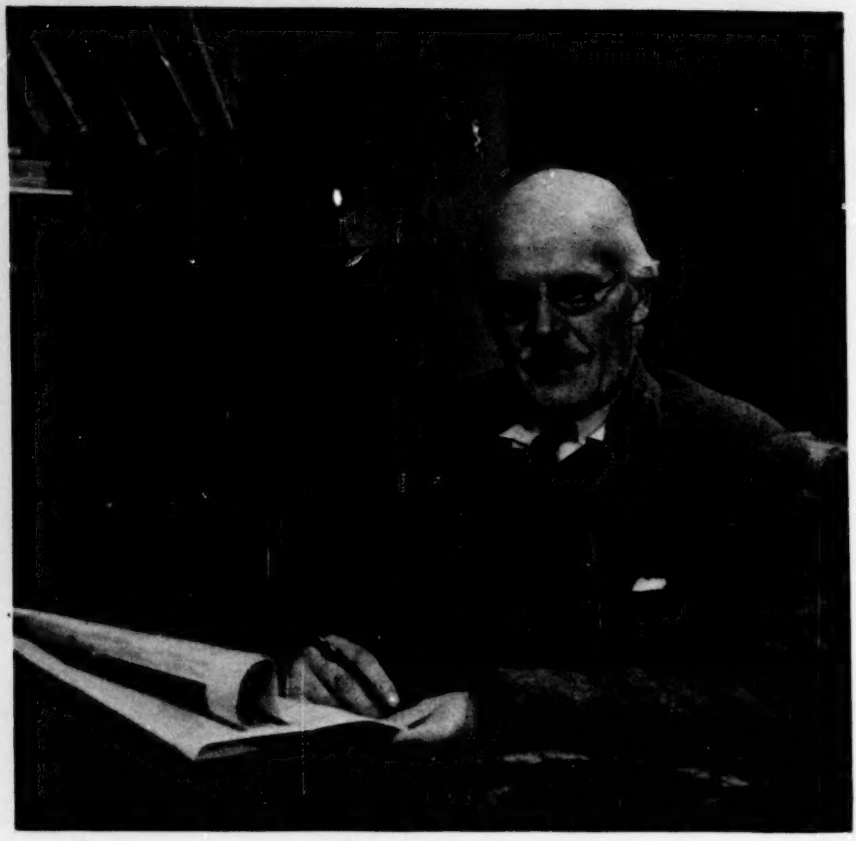
R. L. Wilson, Palaeomagnetism in Northern Ireland.

Part I, The thermal demagnetization of natural magnetic moments in rocks.

Part II, On the reality of a reversal of the Earth's magnetic field.

E. Irving, Palaeomagnetic directions and pole positions. III. Pole numbers 3/1 to 3/87.

GEOPHYSICAL JOURNAL
Vol. 4, 1961



SIR HAROLD JEFFREYS

Copley Medal and Royal Medal of the Royal Society
Gold Medal of the Royal Astronomical Society
Past President of the Royal Astronomical Society

The Earth Today

The knowledge that we have of the Earth today has been gathered almost entirely in the seventy years that we celebrate by the publication of this volume. There was indeed, some idea of the shape of the Earth and of the variation of gravity over it, and the broad outlines of the theory of the figure of the Earth had been worked out in the eighteenth century by Clairaut and others. Local studies, mainly in India, had suggested the idea of isostasy to Sir George Airy and to Archdeacon Pratt, while Sir George Stokes had shown how the detailed shape of the Earth might be found from the variation of gravity over it. But the theory was far in advance of observation for gravity measurements were slow and difficult to make, so that the only results were sparse and unreliable. The study of the interior of the Earth by means of earthquake waves was in its infancy, while with the discovery of radioactivity still five years away, ideas on the age of the Earth and on its thermal history lacked the basis of essential facts. Although the surface of the Earth had been explored geographically, very little was known about the topography of the sea floor and virtually nothing was known about that two-thirds of the Earth's surface that is hidden by the oceans. Mining activities in search of coal and minerals had revealed vertical rock sections several thousand feet thick but the oil industry had not developed the drilling rigs that today can probe deep into the crust.

These various methods of studying the solid Earth were discussed by Sir Harold in ***The Earth***, the first edition of which appeared in 1924. Now greatly expanded, revised and in its fourth edition, this book has guided many in their studies and many there must be who, having read it, regard themselves as Sir Harold's students although they were never at Cambridge nor had the stimulus of his teaching and company. The methods that he then advocated remain the most effective for the study of the Earth: the formulation of rigorous theories and the analysis of data in the light of them by powerful statistical methods where the precision of the data permits, and where the data are less certain, the use of order of magnitude arguments to discriminate between hypotheses. The methods remain effective but the facts to which they are applied are sometimes scarcely recognizable in *The Earth* of 1924. In many respects the development of geophysics since that time has been in the devising of more powerful means of observation, so that mathematical and statistical analysis has been extended to ever more accurate and numerous data. The fourth edition of ***The Earth*** shows the great progress that has been made on these lines, much of it by Sir Harold himself. His contributions to geophysics are assessed in another article, by Dr Stoneley; here let us recall a few developments since 1924. The theory of the figure of the Earth had been completed by Darwin before then but it was not until 1948 that sufficient gravity measurements were available for Sir Harold to undertake a comprehensive reduction to calculate the Earth's external gravity field. This itself is now out of date both on account of the new observations from artificial satellites and of the great increase in the number of gravity measurements at sea. Oldham had discovered the core

of the Earth before 1924 but knowledge of the elastic properties from the surface to the centre was very sketchy by comparison with the detailed results now available. Again the theory of surface waves of both Rayleigh and Love types was known but no applications had been made.

By contrast, there is one subject that stands now in much the same state as in 1924. Now, as then, it does not seem possible to identify the source of the energy that maintains the 14-monthly period in the variation of latitude, as the paper by Professor Munk and Dr Hassan in this volume shows. In a number of subjects, however, there has been a great advance of knowledge but no corresponding body of firm numerical data about the Earth. To mention one of the most striking differences between the Earth of today and of 1924, it seemed then that the general way in which the Earth had come into existence was understood; no one would claim this today. Our understanding of the thermal state of the Earth is related to the question of its origin and here also, although much has been learnt about the processes of heat transfer and the sources of energy in the Earth, yet, as papers in this volume show, many of the questions are without definite answers. Again the processes that have produced the surface features of the Earth are quite obscure and even in the very exact study of seismology there are many unresolved points, particularly concerning the inner core (unsuspected in 1924) and the mantle. The recent confirmation of the existence of a low-velocity layer in the upper part of the mantle, and detailed measurements of the free modes of vibration of the Earth, show that the study of the mechanical properties of the interior of the Earth is still progressing. The old geological theories of continental drift, which have been out of fashion among geophysicists for many years, have come into question again as a result of new experimental evidence from magnetic measurements on ancient rocks and in consequence the mechanisms of movement of rock masses are being investigated in terms of plastic behaviour and of small modifications of the Hookean laws of elasticity.

The papers in this volume represent, we think, some of the most advanced work in geophysics at the present time, especially in the fields in which Sir Harold has been interested. At the same time they demonstrate both the advances and the unsolved problems of the Earth today. In the regions accessible to direct observation, new techniques have brought new assurance: artificial satellites in the study of the gravitational field, instruments of very long period in the study of surface waves and the free oscillations of the Earth, new ways of studying heat flow through the surface and so on. But in the characteristic problems of geophysics as a science, those that involve historical development, the more one knows the less certain one seems to be of the course of events, and this too is shown in some of the contributions in this volume.

Many exciting things have happened in geophysics in recent years. We are glad that Sir Harold has seen with us the launching of artificial satellites and we look forward to still enjoying his company when the exciting things just round the corner take place, when a sample of the mantle is brought up from a deep borehole, for instance, or when the elastic properties of silicates are studied in the laboratory at pressures of a million atmospheres.

SIR HAROLD JEFFREYS

AN APPRECIATION

"The bowels of the harmless earth"—*Shakespeare, King Henry IV, Part I*

To express an appreciation of the contribution of Sir Harold Jeffreys to the science of geophysics is both a privilege and a pleasure; none the less, to present a properly balanced picture is by no means a light task, and the more one knows of the quality and the quantity of his output of research the more difficult is it to make a selection. His high achievement has to be seen against the background of the comparatively slow development of geophysics before World War I, the period at which he began his researches. At that time geophysics could be regarded as a collection of rather miscellaneous investigations covering a wide range of disciplines: it remained for Jeffreys to take these scattered researches and to combine them, along with his own outstanding results, into a connected whole. So closely has he achieved this interconnection of the various topics relating to the solid Earth that one is apt these days to take a short-range view and forget how much the standard expositions owe to his clear insight.

A remarkable combination of astronomical, geodetic and geophysical results was exhibited in Jeffreys's very first paper, *Certain Hypotheses concerning the Internal Structures of the Earth and Moon* published as a Memoir by the R.A.S. In those days a research student had not the organized routine of the present time, with a topic provided by a research supervisor and with regular guidance and discussion. In fact, Jeffreys has said that the germ of the idea in this paper was a remark in one of the dynamics lectures of T. J. I. A. Bromwich: for the rest, Jeffreys did his own foraging. A number of the early papers showed the influence of his training in dynamical astronomy, for instance, his demonstration of the thinness of Saturn's rings: the planets have always been a subject of special interest to him, and older astronomers will remember the consternation of the orthodox when he announced at a meeting of the R.A.S. that the surfaces of the great planets must be very cold. In later times he has returned to the study of the constants of the Solar System. While never forgetting that the Earth is a member of the Solar System, his thoughts were always turning earthwards, and he early seized on the great fund of information that earthquakes provide.

During the First World War, when he worked at the Meteorological Office with Sir Napier Shaw, he became interested in dynamical meteorology, in which he made notable advances. His classification of winds, for example, and his application of tidal theory to monsoons and the general circulation, have been incorporated into the regular text-books of the subject. Somewhat related to these studies are his researches into tidal theory, the oscillations of lakes and the generation of water waves by wind. His interest in the problem of a liquid heated below (called by him the "porridge problem") has led to a number of papers by himself and others on a topic of wide application.

It is, perhaps, his long series of researches on the interior of the Earth, its temperature and constitution, that have been the greatest stimulus to others. His close acquaintance with the classical investigations of Stokes, Rayleigh, Love and George Darwin have enabled him to advance on a broad front, keeping in mind always that the older work was usually a control on newer theories. One link, for example, between the old and the new was provided by his demonstration that the solid "core" of the Wiechert hypothesis could be identified with the fluid core of Oldham and Gutenberg by allowance for compressibility, without introducing *ad hoc* hypotheses.

It does not suffice, however, to have a fairly detailed knowledge of the interior of the Earth as it is at the present day; that information should be co-ordinated with the past history, and so with cosmology. The origin of the Solar System is an old problem, and is still with us; Jeffreys took important steps towards discriminating among the various theories propounded. The age of the Earth is an important datum in any discussion of the development of the outer layers, and the treatment by Holmes and by Jeffreys of the heat conduction problem, in which the data are the conductivities of the rocks and the rate of generation of heat by radioactivity together with the surface gradient of temperature, led to an age of about 3×10^9 years since the formation of the crust, and to a layering of the continents in reasonable agreement with the indications of seismology.

To attempt to summarize the achievements of Jeffreys in seismology would be beyond the bounds of this article. He early appreciated the importance of H. H. Turner's early attempts to improve the tables of travel-times of *P* and *S*, and his 1931 tables, followed by the *J–B* tables compiled jointly with K. E. Bullen in 1939, have by successive approximation produced the tables now in general use. They have provided a powerful instrument of research: the well-known technique of the calculus of variations enables the wave-velocity to be calculated at all depths reached by the waves, and Jeffreys and Bullen gathered a rich harvest of information, so that except for the Lehmann "inner core", the distribution of density and elastic properties may be considered well known. Jeffreys has published a number of important papers on the theory of surface elastic waves, and over a long period has been intensely interested in the formulation of the laws of imperfection of elasticity. His discussions of new earthquakes in Europe greatly stimulated studies of the layering of the crust of the continents, and it is only in later days that seismic exploration at sea has given analogous information about the constitution of the ocean floor. In view of the attention now being paid to "Project Moho", his words in the first edition of *The Earth* in 1924 were prophetic: ". . . up to the present no mine or boring has been made in the ocean bottom. If such an undertaking could be carried out, most valuable information could be obtained."

The theory of the figure of the Earth has been relevant to many of Jeffreys's researches, including his first paper, mentioned above. Avoiding the over-simplifications of some enthusiastic isostasists, he has greatly helped our understanding of the mechanism of the support of surface inequalities and the development of surface features. In later years he has returned to the expression of the Earth's gravitational field by spherical harmonic terms, a topic of present-day interest to those engaged in discussing the orbits of artificial satellites.

At least one other aspect of his work must be stressed. While astronomers (trained in "Combination of Observations") have all along been accustomed to indicate the precision of their results, it has to be admitted that many earlier geophysicists were lax in this respect, and tended to place too great a reliance on

graphical methods; Jeffreys, who was conducting researches in the theory of probability contemporaneously with his development of seismology, was responsible for systematic insistence on the use of a statistical criterion in appraising geophysical results. Least squares solutions replaced graphical methods, discussions of frequency distributions superseded crude averages, and tests of significance were regularly employed. The history of this change may be discerned in the geophysical examples used as illustrations in Jeffreys's *Theory of Probability*.

A word about his books. His first book, *The Earth*, now in its (much-developed) fourth edition, was destined to become the geophysicists' bible. His *Scientific Inference*, in two editions, and the two editions of the *Theory of Probability* provide a course of mental discipline that, one fears, is still badly needed by physical scientists. *Earthquakes and Mountains* conveys in non-mathematical form much geophysics of the solid Earth. Professor Jeffreys made a valuable contribution to mathematical technique in his *Operational Methods in Mathematical Physics*; his *Cartesian Tensors* had a marked impact on the employment of tensor notation. Of enormous benefit to research workers is the large volume *Mathematical Physics*, written jointly with his wife, Lady Jeffreys; it is all the more valuable because the topics are expounded in relation to the everyday needs of scientists.

In contributing to the advancement of geophysics he has been unsparing in giving advice, help and encouragement to other investigators. Mention must be made of his long period of service as Editor of the *Geophysical Supplement*, in which he succeeded H. H. Turner; the quality and prestige of the *Supplement* were much enhanced by his many published contributions. Following H. H. Turner and H. H. Plaskett he became Honorary Director of the International Seismological Summary at the time it came formally under the aegis of the International Association of Seismology in 1948, when it was agreed by that Association that the Jeffreys–Bullen tables should henceforth be used for the routine determination of epicentres. It was fitting that he should have been chosen President of the International Association of Seismology and the Physics of the Earth's Interior for the triennium 1957–1960, which included the International Geophysical Year.

The work of Harold Jeffreys has received recognition by many Academies, Learned Societies and Universities. It is pleasing to record that "as we go to press" the Royal Society has bestowed on him its highest award, the Copley Medal. Long may he live to enjoy the years of retirement and to enhance the prestige of geophysics!

R. STONELEY

I. The Earth's Gravitational Field and Geodesy

PAGE

D. G. KING-HELE, The Earth's gravitational potential, deduced from the orbits of artificial satellites 3

R. H. MERSON, The motion of a satellite in an axi-symmetric gravitational field 17

A. H. COOK, Resonant orbits of artificial satellites and longitude terms in the Earth's external gravitational potential 53

C. TSUBOI, Numerical tables useful for studying the gravitational field at higher elevations of the Earth 73

A. MARUSSI, Intrinsic coordinates in practical geodesy 83

The Earth's Gravitational Potential, deduced from the Orbits of Artificial Satellites

D. G. King-Hele

"And measure every wand'ring planet's course,
Still climbing after knowledge infinite."—*C. Marlowe,*
Tamburlaine the Great, Pt. I, II, 7.

Summary

The numerical values of several of the higher harmonics in the Earth's gravitational potential have been obtained from the perturbations to the orbits of satellites. This paper describes the relevant methods, cursorily reviews previous results, and gives details of a fresh determination of the second, fourth and sixth harmonics from the orbits of Sputnik 2, Vanguard 1 and Explorer 7.

1. Introduction

Artificial satellites offer a powerful method of investigating the Earth's gravitational field, since several of the orbital perturbations caused by the non-spherical components of the Earth's field can be measured very accurately. The orbit of a satellite *in vacuo* about an isolated, spherical Earth would be an exact ellipse of fixed size and shape, lying in a plane fixed in direction relative to the stars, and the orientation of the ellipse in its own plane would not change. This ellipse is perturbed in various ways when the higher harmonics in the Earth's gravitational potential are taken into account. The most important perturbations are two secular changes, a steady rotation of the orbital plane about the Earth's axis (regression of the nodes), and a rotation of the major axis of the orbit in its own plane. From these rotations, which can be accurately measured, several of the even harmonics in the Earth's potential have already been evaluated. In addition there are short-period orbital variations in the course of one revolution, which have so far proved too small to measure, and longer-term oscillations whose period is of the same order as the period of rotation of the major axis (at least 20 days). These latter oscillations have indicated the magnitude of the third and fifth harmonics.

It is fortunate that, to a first approximation, these gravitational perturbations to the orbit do not become confused with other perturbations, such as those due to the atmosphere and the Moon, which are either different in kind or very much smaller. For example, the main effect of the atmosphere is to make the orbit contract and become more circular; its effect on the rotation of the orbital plane is almost negligible. In an accurate analysis a few of these minor perturbations have to be taken into account, but since they are so small, their values do not need to be known very precisely; the corrections can be made without introducing significant error.

In this paper the methods for evaluating the later harmonics from satellite orbits are sketched, previous results are summarized, and a new determination of the second, fourth and sixth harmonics is made from the orbits of Sputnik 2, Vanguard 1 and Explorer 7.

2. Gravitational potential

The appropriate general form for the gravitational potential U due to a nearly spherical body such as the Earth, at an exterior point distant r from the centre, is an expansion in terms of spherical harmonics $S_n(\theta, \phi)$:

$$U = \sum_{n=0}^{\infty} A_n r^{-n-1} S_n(\theta, \phi) \tag{1}$$

where θ is taken as the angle between the radius vector and the Earth's axis (the co-latitude), ϕ is the longitude and the A_n are constants.

As a result of the Earth's rotation, satellites tend to sample all longitudes impartially, and it is therefore unlikely that they will reveal the variation of the potential with longitude. (The only exceptions to this rule (A. H. Cook 1960) are satellites with periods near 24 hours, but so far there have been none of these.) Consequently it is, for the present, permissible to write the potential as if it were independent of longitude. Equation (1) may then be expressed more explicitly in terms of P_n, the Legendre polynomial of order n, as

$$U = \frac{GM}{R}\left\{\frac{R}{r} - \sum_{n=2}^{\infty} J_n \left(\frac{R}{r}\right)^{n+1} P_n(\cos\theta)\right\} \tag{2}$$

where G is the gravitational constant, M the mass of the Earth, and R the Earth's equatorial radius. ($R = 6378{\cdot}1 \pm 0{\cdot}1$ km, and $GM/R = 62{\cdot}494 \pm 0{\cdot}001$ km^2/s^2.) The J_n are constants whose values are to be determined, and J_1 is zero if the equatorial plane is chosen to pass through the Earth's centre of mass. The explicit forms of the first five relevant Legendre polynomials are:

$$\begin{aligned}
P_2 &= \tfrac{1}{2}(3\cos^2\theta - 1)\\
P_3 &= \tfrac{1}{2}(5\cos^3\theta - 3\cos\theta)\\
P_4 &= \tfrac{1}{8}(35\cos^4\theta - 30\cos^2\theta + 3)\\
P_5 &= \tfrac{1}{8}(63\cos^5\theta - 70\cos^3\theta + 15\cos\theta)\\
P_6 &= \tfrac{1}{16}(231\cos^6\theta - 315\cos^4\theta + 105\cos^2\theta - 5).
\end{aligned}$$

The more familiar form for the gravitational potential, in terms of the second and fourth harmonics only (Jeffreys 1952),

$$U = \frac{GM}{R}\left\{\frac{R}{r} + J\frac{R^3}{r^3}(\tfrac{1}{3} - \cos^2\theta) + \frac{D}{35}\frac{R^5}{r^5}(35\cos^4\theta - 30\cos^2\theta + 3)\right\} \tag{3}$$

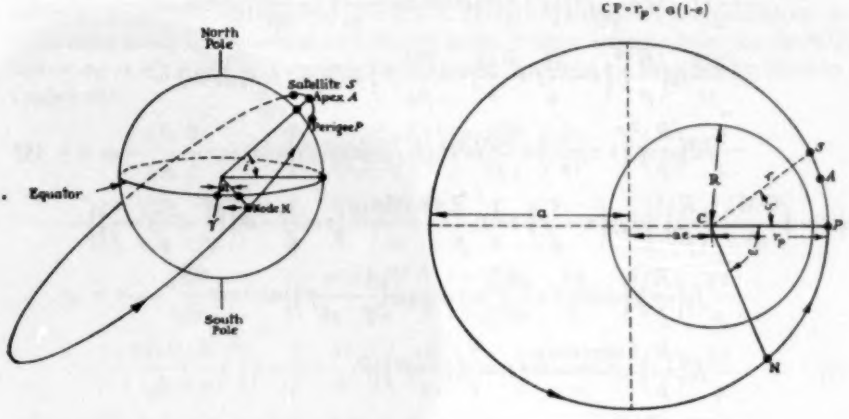
is not sufficiently general now that the coefficients of further harmonics can be evaluated, and equation (2) seems the most appropriate generalization of it, the choice of the symbol J being intended as a tacit tribute to Sir Harold Jeffreys. The coefficients J and D are related to J_2 and J_4 as follows:

$$J = \frac{3}{2}J_2, \qquad D = -\frac{35}{8}J_4.$$

It will be recalled that, before the advent of satellites, only the coefficient of the second harmonic had been adequately determined, the most generally accepted value (Jeffreys 1952) being $J = 1\,637 \times 10^{-6}$ ($\therefore J_2 = 1\,091 \times 10^{-6}$). The coefficient D was given a conventional value of $10{\cdot}6 \times 10^{-6}$, chosen to make the meridional section of the geoid an exact ellipse.

3. Method of finding the J_n from satellite orbits

The equations of motion of a satellite moving in the gravitational field specified by (2) can readily be written down, and, not so readily, integrated to give the orbital perturbations as functions of the J_n and of the orbital elements, namely a the semi major axis, e the eccentricity, i the inclination of the orbit to the equator, and ω the argument of perigee (see Figure 1 for diagrammatic definitions). Thus,



(*a*) Relative to Earth. (*b*) In its own plane.

FIG. 1. Diagram of Satellite orbit.

in principle, any perturbation P may be expressed in terms of functions F of the elements, as

$$P = \sum_{n=2}^{\infty} J_n F_n(a,e,i,\omega) + \sum_{n=2}^{\infty}\sum_{m=2}^{\infty} J_n J_m F_{nm}(a,e,i,\omega) + \dots . \qquad (4)$$

Since the J_n are of order 10^{-6} for $n > 2$, it has so far only been necessary to evaluate the terms in the first series, together with the J_2^2 term in the second series.

The orbital theory from which equations like (4) can be derived has now been developed by various authors (Krause 1956; Roberson 1957; King-Hele 1958; Kozai 1959c; Brouwer 1959; Groves 1960; Zhongolovitch 1960; Merson 1960) using different methods, and there seems to be general agreement between them on the form of functions F_n in the first series in equation (4). The most generalized version of these results is that of Groves (1960), and the most complete explicit forms are those obtained by Merson (1960), whose expressions for the west–east rate of rotation of the orbital plane, $\dot{\Omega}$ (i.e. the rate of change of the right ascension of the ascending node), and the rate of rotation of the major axis of the ellipse, $\dot{\omega}$ (i.e. the rate of change of the argument of perigee), are given below, shorn of a few

negligibly small terms, and with the addition of the J_2^2 term given by King-Hele (1958).

$$\dot{\Omega} = -\left(\frac{GM}{a^3}\right)^{\frac{1}{2}}\left(\frac{R}{p}\right)^2\cos i\left[\frac{3}{2}J_2-\frac{3}{2}J_3\frac{R}{p}\frac{e\sin\omega}{\sin i}\left(1-\frac{15}{4}f\right)-\right.$$
$$-\frac{15}{4}J_4\left(\frac{R}{p}\right)^2\left\{\left(1-\frac{7}{4}f\right)\left(1+\frac{3}{2}e^2\right)+\left(\frac{7}{4}f-\frac{3}{4}\right)e^2\cos 2\omega\right\}+$$
$$+\frac{15}{4}J_5\left(\frac{R}{p}\right)^3 e\operatorname{cosec} i\left\{\left(1-\frac{21}{2}f+\frac{105}{8}f^2\right)\left(1+\frac{3}{4}e^2\right)\sin\omega+\right.$$
$$\left.+\frac{7}{8}\left(1-\frac{15}{8}f\right)e^2\sin 3\omega\right\}+\frac{105}{16}J_6\left(\frac{R}{p}\right)^4\left\{\left(1-\frac{9}{2}f+\frac{33}{8}f^2\right)\left(1+5e^2+\frac{15}{8}e^4\right)-\right.$$
$$\left.-\frac{5}{2}\left(1-6f+\frac{99}{16}f^2\right)\left(1+\frac{1}{2}e^2\right)e^2\cos 2\omega-\frac{15}{32}\left(1-\frac{33}{20}f\right)fe^4\cos 4\omega\right\}-$$
$$-\frac{315}{32}J_8\left(\frac{R}{p}\right)^6\left(1-\frac{33}{4}f+\frac{143}{8}f^2-\frac{715}{64}f^3\right)+O(J_7, J_8e^2, J_9\ldots)-$$
$$\left.-\frac{9}{4}J_2^2\left(\frac{R}{p}\right)^2\left(1-\frac{19}{12}f\right)+O(J_2^2e^2, J_2J_3, J_2J_4\ldots)\right]. \tag{5}$$

$$\dot{\omega} = \left(\frac{GM}{a^3}\right)^{\frac{1}{2}}\left(\frac{R}{p}\right)^2\left[3J_2\left(1-\frac{5}{4}f\right)+\frac{3}{2}J_3\frac{R}{p}\frac{\sin i\sin\omega}{e}\left\{1-\frac{5}{4}f+e^2\left(\frac{35}{4}-\frac{35}{4}f-\frac{1}{f}\right)\right\}-\right.$$
$$-\frac{15}{2}J_4\left(\frac{R}{p}\right)^2\left\{1-\frac{31}{8}f+\frac{49}{16}f^2+f\cos 2\omega\left(\frac{3}{8}-\frac{7}{16}f\right)\right\}-$$
$$-\frac{15}{4}J_5\left(\frac{R}{p}\right)^3\frac{\sin i\sin\omega}{e}\left(1-\frac{7}{2}f+\frac{21}{8}f^2\right)+$$
$$+\frac{105}{8}J_6\left(\frac{R}{p}\right)^4\left\{\left(1-8f+\frac{129}{8}f^2-\frac{297}{32}f^3\right)+\right.$$
$$\left.+\left(\frac{5}{4}-\frac{15}{4}f+\frac{165}{64}f^2\right)f\cos 2\omega\right\}-$$
$$-\frac{315}{4}J_8\left(\frac{R}{p}\right)^6\left(1-\frac{331}{32}f+\frac{2057}{64}f^2-\frac{19877}{512}f^3+\frac{16445}{1024}f^4\right)+$$
$$+O(J_4e^2, J_5e^2, J_6e^2, J_7, J_8\cos 2\omega, J_8e^2, J_9, J_{10}\ldots)+$$
$$\left.+O(J_2^2, J_2J_3, \ldots)\right] \tag{6}$$

where $p = a(1-e^2)$, $f = \sin^2 i$. The value of a can be found from the nodal period T_N by means of the equation

$$T_N = 2\pi\left(\frac{a^3}{GM}\right)^{\frac{1}{2}}\left\{1-\frac{3J_2}{8}\left(\frac{R}{p}\right)^2(7\cos^2 i-1)+O(J_2e)\right\}.$$

In these equations the orbital elements, a, e and i, are to be taken not as osculating elements but as certain mean elements. The need for an exact definition of the elements only arises in the J_2^2 term, whose form is completely different if, for

example, osculating elements evaluated at the ascending node are used. The J_2^2 term in $\dot{\Omega}$ given above, that derived by King-Hele (1958), has subsequently been verified under more stringent assumptions by Message (1960) and Merson (1960), who discusses the question in more detail.

It will be noted that some of the terms in equations (5) and (6) are not strictly "secular", since they contain the periodic factors $\cos\omega$, $\cos 2\omega$, etc. It is desirable to include them in the same formulae, however, partly for the sake of completeness and partly because ω may change by only a small fraction of a cycle during the lifetime of a satellite, thus allowing the long-period terms to masquerade successfully as secular terms.

Of the other orbital elements, the semi major axis a remains constant, apart perhaps from a term of order $J_2^2 e$, primarily because it represents the constant in the equation for conservation of energy, while the eccentricity and inclination, e and i, undergo no secular variation (if "secular" is interpreted strictly), but do exhibit long-period oscillations. Since a is effectively constant, the oscillation in e can most easily be visualized as a change in the perigee distance from the Earth's centre $r_p = a(1-e)$. The appropriate formulae for i and r_p, as given by Merson (1960) are:

$$i = i_0 + \frac{eJ_3}{2J_2}\frac{R}{p}\cos i \sin\omega + \frac{15J_4e^2}{64J_2}\left(\frac{R}{p}\right)^2 \frac{1-7f/6}{1-5f/4}\sin 2i(1-\cos 2\omega) - \\ - \frac{5eJ_5}{4J_2}\left(\frac{R}{p}\right)^3\left(1-\frac{7}{2}f+\frac{21}{8}f^2\right)\left(1-\frac{5}{4}f\right)^{-1}\cos i\sin\omega + O(J_6e^2, J_2^2e) \qquad (7)$$

$$r_p = r_{p0} + \frac{J_3R}{2J_2}\sin i \sin\omega + \frac{15eJ_4Rf}{32J_2}\frac{R}{p}\frac{1-7f/6}{1-5f/4}(1-\cos 2\omega) - \\ - \frac{5J_5R}{4J_2}\left(\frac{R}{p}\right)^2\left(1-\frac{7}{2}f+\frac{21}{8}f^2\right)\left(1-\frac{5}{4}f\right)^{-1}\sin i\sin\omega + O(J_6e, J_2^2) \qquad (8)$$

when suffix o denotes values when $\omega = 0$.

Equations (5)–(8) can be utilized in various ways. Ideally, perhaps, all the equations should be applied to all observed orbits, with suitable weighting according to the accuracy of the orbital determinations; from the resulting mass of equations a set of "best values" of the J_n could be obtained. In practice the equations have, so far, been used in a piecemeal fashion, and applied only to the best-determined orbits: but the results are probably almost as good as would be obtained from the ideal process.

For an orbit with a rapidly rotating major axis, the periodic terms in equations (5)–(8) yield equations for J_3 and J_5. (The contributions from J_4 can be excluded, since they have a different period and moreover are small as a result of the factor e.) For example, equation (8) shows that the distance of perigee from the Earth's centre oscillates with the same period as ω and amplitude

$$\frac{R}{2}\sin i\left\{\frac{J_3}{J_2} - \frac{5}{2}\frac{J_5}{J_2}\left(\frac{R}{p}\right)^2\left(1-\frac{7}{2}f+\frac{21}{8}f^2\right)\left(1-\frac{5f}{4}\right)^{-1}\right\}. \qquad (9)$$

Thus the observed amplitude of oscillation of each orbital element yields an equation J_3, J_5, ... from which the values of the odd coefficients can be found. In practice the amplitude of the oscillation in perigee distance, (9), is the most accurately observed of these quantities, but the oscillations in i, ω and Ω are also

measurable. It therefore proves possible to obtain a value of J_3 with quite good accuracy, together with a tentative value of J_5, if $J_7, J_9, \ldots$ are ignored.

The strictly secular terms in equations (5) and (6) can be utilized to evaluate the even J_n. The contributions of the odd J_n to equations (5) and (6) are small, because of the existence of multiplying factors of the form $e\sin\omega$, etc., which are of order 0·1. If the odd-J_n terms are still appreciable, their influence can be further reduced by choosing a time interval during which the mean value of $\sin\omega$ is small. Consequently, the J_3 and J_5 terms in (5) and (6) can often be ignored; and, if not, the imperfectly-known values of J_3 and J_5 can be substituted without introducing appreciable errors. Thus equation (5), with an observed value of $\dot{\Omega}$ on the left-hand side, and an approximate value for J_2^2, provides in effect one linear relation between the even J_n. If there are r observed values of $\dot{\Omega}$ available, we have r equations between the J_n, from which the values of $J_2, J_4, \ldots J_{2r}$ can in principle be determined, if J_{2r+2}, J_{2r+4} ... are assumed zero. In practice, the observed values must be derived from satellites with appreciably different values of either the inclination i or the semi-latus rectum p: otherwise the simultaneous equations will tend to be ill-conditioned. Similar conclusions apply to equations (6). If, however, a total of r values of $\dot{\Omega}$ and $\dot{\omega}$ is available from suitably chosen satellites, it should be possible to evaluate $J_2, J_4 \ldots J_{2r}$.

To sum up, therefore, equations (7) and (8) and the periodic terms in (5) and (6) serve to determine the odd J_n, while the secular terms in equations (5) and (6) give the even J_n. When accurate satellite orbits become available as a matter of routine, it should be possible to evaluate a large number of the J_n.

The observed values for use in these equations must of course be purged of any contributions from other sources of perturbation. Three such sources which deserve notice are the effects of the atmosphere, the Sun and the Moon. G. E. Cook (1960) has shown that for an oblate, rotating atmosphere of ellipticity ϵ, in which air density ρ varies exponentially with height y, so that $\rho \propto \exp(-\beta y)$, the changes in Ω and ω due to aerodynamic forces, $\Delta\Omega_A$ and $\Delta\omega_A$, say, may be expressed in terms of the change Δa in the semi major axis by the equations

$$\left.\begin{aligned}\frac{\Delta\Omega_A}{\Delta a} &= \frac{I_2-2eI_1+cI_4\cos 2\omega}{I_0+2eI_1+cI_2\cos 2\omega}\cdot\frac{T_d\,\Lambda\sin 2\omega}{4a(1-T_d\,\Lambda\cos i)}\{1+O(e^2)\}\\ \frac{\Delta\omega_A}{\Delta a} &= -\cos i\frac{\Delta\Omega_A}{\Delta a}-\frac{2cI_2-\tfrac{1}{2}ce(I_1-9I_3)+2c^2I_4\cos 2\omega}{I_0+2eI_1+cI_2\cos 2\omega}\cdot\frac{\sin 2\omega}{\beta a^2e^2}\{1+O(e^2)\}\end{aligned}\right\} \quad (10)$$

where $c = \frac{1}{2}\beta\epsilon a(1-\epsilon)\sin^2 i$, T_d = orbital period, expressed as a fraction of a sidereal day, Λ = angular velocity of the atmosphere divided by the angular velocity of Earth, and I_n is the Bessel function of the first kind and imaginary argument, of order n and argument βae. Values of $\dot{\Omega}_A$ and $\dot{\omega}_A$ for a particular satellite can be found on multiplying each of equations (10) by the known values of da/dt for that satellite.

Kozai (1959b) has given the contributions to $\dot{\Omega}$ and $\dot{\omega}$ resulting from the Moon's attracton, $\Delta\Omega_M$ and $\Delta\omega_M$ say, in the form:

$$\left.\begin{aligned}\Delta\dot{\Omega}_M &= -\frac{3}{4}\frac{n_M^2}{n}m_M\left(1-\frac{3}{2}b^2\right)(1-e^2)^{-\frac{1}{2}}\left(1+\frac{3}{2}e^2\right)\cos i\\ \Delta\dot{\omega}_M &= \frac{3}{4}\frac{n_M^2}{n}m_M\left(1-\frac{3}{2}b^2\right)(1-e^2)^{-\frac{1}{2}}\left(2-\frac{5}{2}\sin^2 i+\frac{e^2}{2}\right)\end{aligned}\right\} \quad (11)$$

where n_M is the angular velocity of the Moon, n that of the satellite, m_M denotes the ratio of the Moon's mass to the Earth's, and

$$b^2 = \tfrac{1}{2}\sin^2 J'(1+\cos^2\epsilon') + \sin^2\epsilon' \cos^2 J' \qquad (12)$$

ϵ' being the obliquity of the ecliptic, and J' the inclination of the Moon's orbit to the ecliptic. Values for the solar perturbations, suffix s, which are about half those due to the Moon, can be obtained from equations (11) by putting $m_M = 1$, $J' = 0$, and changing n_M to n_s.

Equations (10) and (11) give us values for $\dot{\Omega}_A$, $\dot{\omega}_A$, $\dot{\Omega}_M$, $\dot{\omega}_M$, $\dot{\Omega}_S$ and $\dot{\omega}_S$. These contributions to $\dot{\Omega}$ and $\dot{\omega}$, which are usually of the same order as the observational errors, or smaller, must be subtracted from the observed values of $\dot{\Omega}$ and $\dot{\omega}$ before the latter are substituted into equations (5) and (6). As far as is known, the other perturbations to the orbit, due to relativity effects (Krause 1956, King-Hele 1958) electromagnetic forces, the effects of the planets (King-Hele 1958) and solar radiation pressure (Musen 1960) are negligible, to the order of accuracy so far achieved.

4. Previous determinations of J_n

4.1 *Odd values of n*

The value of J_3 was first determined from the orbit of Vanguard 1: it was found that the perigee distance from the Earth's centre oscillated with a period of about 80 days, almost equal to the period of rotation of the major axis, and an amplitude of about 4 km. O'Keefe, Eckels and Squires (1959a), applying equation (8) with J_4 and J_5 zero, deduced that $J_3 = (-2{\cdot}4 \pm 0{\cdot}2) \times 10^{-6}$. An independent analysis by Kozai (1959a), using also the variations in i, ω and Ω, gave $J_3 = (-2{\cdot}2 \pm 0{\cdot}1) \times 10^{-6}$. A subsequent more detailed study along similar lines by O'Keefe, Eckels and Squires (1959b) led to the values

$$\left.\begin{aligned} J_3 &= (-2{\cdot}4 \pm 0{\cdot}3) \times 10^{-6} \\ J_5 &= (-0{\cdot}1 \pm 0{\cdot}1) \times 10^{-6} \end{aligned}\right\} \qquad (13)$$

and at the time of writing these values stand as the best available for J_3 and J_5, though Sir Harold Jeffreys (private communication) contends that the standard deviations should be smaller. Recently Zhongolovitch (1960b) has succeeded in obtaining an approximate value of J_3, of $(-2 \pm 3) \times 10^{-6}$, from Sputniks 1, 2 and 3, while Cohen & Anderle (1960) have verified the value of J_3 with the aid of the Transit 1B satellite.

4.2. *Even values of n*

In the first six months of 1958 it was noted that the values of $\dot{\Omega}$ obtained from the satellites Sputnik 2, Explorer 1, and Vanguard 1 were all rather lower than would be expected if J_2 had the value $1\,091 \times 10^{-6}$, and various authors (Buchar 1958, Harvard 1958, Merson & King-Hele 1958) derived values of J_2, ranging between $1\,082 \times 10^{-6}$ and $1\,086 \times 10^{-6}$, by taking J_4 as either zero or a conventional value.

On combining the values of $\dot{\Omega}$ from Sputnik 2 (inclination 65°) and Vanguard 1 ($i = 34°$), Jacchia (1958) obtained $J_2 = 1\,082{\cdot}7 \times 10^{-6}$ and $J_4 = -2 \times 10^{-6}$, and King-Hele (1959a) found $J_2 = 1\,083{\cdot}3 \times 10^{-6}$ and $J_4 = -1 \times 10^{-6}$.

The possibility remained, however, of a substantial contribution from the J_6 term: King-Hele & Merson (1959) obtained a tentative value of J_6 by combining values of $\dot{\Omega}$ from Sputnik 2 ($i = 65°$), Vanguard 1 ($i = 34°$) and Explorer 4 ($i = 50°$) to give $J_2 = (1083{\cdot}1 \pm 0{\cdot}2) \times 10^{-6}$, $J_4 = (-1{\cdot}4 \pm 0{\cdot}2) \times 10^{-6}$ and $J_6 = (0{\cdot}1 \pm 1{\cdot}5) \times 10^{-6}$. Further refinement of the same method, taking account of J_3 and of the perturbations expressed in equations (10)–(12), and using all the relevant terms in (5), gave (King-Hele 1959b)

$$\left.\begin{aligned} J_2 &= (1083{\cdot}0 \pm 0{\cdot}2) \times 10^{-6} \\ J_4 &= (-1{\cdot}3 \pm 0{\cdot}2) \times 10^{-6} \\ J_6 &= (-0{\cdot}1 \pm 1{\cdot}5) \times 10^{-6} \end{aligned}\right\}. \qquad (14)$$

The standard errors quoted are those arising from errors in the observational values: larger errors could arise as a result of ignoring J_8, etc.

Values of J_2 and J_4 have also been calculated by a different method, using $\dot{\Omega}$ and $\dot{\omega}$ from Vanguard 1, and ignoring the J_2^2 and J_6 terms. The first analysis of this type (Lecar, Sorenson & Eckels 1959) gave $J_2 = 1082 \times 10^{-6}$ and $J_4 = -2{\cdot}4 \times 10^{-6}$, and in a subsequent study, O'Keefe, Eckels & Squires (1959b) found

$$\left.\begin{aligned} J_2 &= (1082{\cdot}5 \pm 0{\cdot}1) \times 10^{-6} \\ J_4 &= (-1{\cdot}7 \pm 0{\cdot}1) \times 10^{-6} \end{aligned}\right\} \qquad (15)$$

the standard errors quoted being due to the errors in the observational values. Again larger errors may occur as a result of the neglected J_2^2, J_6, J_8 terms.

The values obtained in equations (14) and (15) are derived from independent observational values, for different satellites. The close agreement between them is highly satisfactory, and suggests that the neglect of later terms in the series expansion of the potential is probably not causing serious errors.

These previous results have been reviewed with fuller reference to their geophysical implications by A. H. Cook (1959) and Bomford (1960).

Recently, Zhongolovitch (1960b), using $\dot{\Omega}$ and $\dot{\omega}$ from Sputnik 2, Sputnik 3 and its rocket, ignoring J_6 but taking account of the J_2^2 terms, has obtained the following values:

$$J_2 = (1083{\cdot}3 \pm 0{\cdot}7) \times 10^{-6}$$
$$J_4 = (-4{\cdot}1 \pm 0{\cdot}7) \times 10^{-6}.$$

His value of J_4 differs significantly from those of equations (14) and (15).

(The results of two further evaluations of the J_n became available after this paper had been written, too late to be commented upon in Section 6. Kozai (1960) found

$$J_2 = (1082{\cdot}19 \pm 0{\cdot}03) \times 10^{-6}, \qquad J_3 = (-2{\cdot}29 \pm 0{\cdot}02) \times 10^{-6},$$
$$J_4 = (-2{\cdot}1 \pm 0{\cdot}1) \times 10^{-6}, \qquad J_5 = (-2{\cdot}3 \pm 0{\cdot}2) \times 10^{-7}.$$

Buchar (1960) found

$$J_2 = (1083 \pm 0{\cdot}4) \times 10^{-6}, \qquad J_4 = (-1{\cdot}1 \pm 0{\cdot}5) \times 10^{-6}.)$$

5. A fresh determination of J_2, J_4 and J_6 from Sputnik 2, Vanguard 1 and Explorer 7

5.1. *Introduction*

Better values of the even J_n can usually be obtained from $\dot{\Omega}$ than from $\dot{\omega}$, because Ω is derived directly from observation, whereas ω can be found only if the rest of the orbital elements, and in particular the eccentricity, are accurately known. It may sometimes happen that for an orbit which has high eccentricity (near 0·2) and is little affected by the atmosphere, such as that of Vanguard 1, $\dot{\omega}$ can be found with adequate accuracy. In general, however, the values of $\dot{\Omega}$ are likely to be more reliable.

If this viewpoint is adopted, we must look for satellites whose orbital inclinations are as widely separated as possible, so that the simultaneous equations obtained from the use of equation (5) do not become ill-conditioned. The orbital inclinations of satellites have so far (May 1960) been confined to four narrow bands, largely because of the layout of the range instrumentation near the launching sites. Satellites launched from Cape Canaveral have had inclinations either between 32° and 34½° (10 satellites in all, including final-stage rockets, etc.) or between 47° and 51½° (9 in all); those launched from Russia (14, excluding Lunik 3) have had inclinations between 64½° and 65½°; those launched from Vandenberg Air Force Base (7) have had inclinations between 79° and 84°, except for Discoverer 2, whose inclination of 90° is, however, useless for our purposes, since $\dot{\Omega} = 0$.

Since the spread of inclination within each of these four groups is less than 5°, it seems advisable not to choose two satellites from the same group but to select the best satellite from each group. In the first group the satellite whose orbit is best determined is Vanguard 1. In the second group, the five satellites resulting from the Tiros and Transit launchings in April 1960 must be excluded because they have not yet been in orbit long enough, and Explorer 6 because it was lost before the end of 1959. That leaves Explorer 4, Explorer 7 and its rocket: of these, Explorer 7 is to be preferred, since it has been well observed, both optically and by radio, and its orbit is little affected by the atmosphere. In the third group, the Russian satellites, Sputnik 2 has been chosen, since it has the best-established orbit. The fourth group consists of the Discoverer satellites, and unfortunately none of these has had a long enough lifetime to be entirely suitable for use in determining the J_n. The best results would be likely to come from Discoverer 6 or Discoverer 5 capsule.

5.2. *Orbital data*

5.2.1. *Sputnik* 2.—The orbit of Sputnik 2 was determined by R. H. Merson from kinetheodolite observations made in Britain and theodolite observations at Jokioinen, Finland. Details have been given by King-Hele & Merson (1958) and Cornford (1958). It was found (Merson & King-Hele 1958) that the observed value of $\dot{\Omega}$ was less during the whole lifetime, by a factor $1{\cdot}0068 \pm 0{\cdot}0003$ (standard deviation), than the value obtained by substituting in equation (5) $J_2 = \frac{2}{3} \times 1\,637 \times 10^{-6}$, $J_4 = -\frac{3}{35} \times 10{\cdot}6 \times 10^{-6}$, $J_3 = J_5 = J_7 = J_8 = 0$ and ignoring the $J_4 e^2$ terms. This observed value of $\dot{\Omega}$ has been corrected for the effects of atmospheric rotation and luni-solar perturbations, which contribute respectively 1/10 000 and 1/20 000 of the total $\dot{\Omega}$, to give a corrected "observed value" for use on the left-hand side of equation (5). The terms within the square brackets on the right-hand side of (5)

have been evaluated for 1958 March 14, when $i = 65{\cdot}24°$, $e = 0{\cdot}048$, $R/p = 0{\cdot}9258$, except for the J_3 term, in which it was thought advisable to use the average value of $e\sin\omega$ during the lifetime of the satellite, 0·037, rather than the value on March 14, which happens to be exceptionally small. In the J_2^2 term in (5), J_2 was taken as 1083×10^{-6}. If J_8, J_{10} ... are assumed zero, the resulting equation between J_2, J_4 and J_6 is

$$J_2+0{\cdot}949J_4+0{\cdot}303J_6 = 1081{\cdot}71(1\pm0{\cdot}0003)\times10^{-6}\text{—}\textit{Sputnik}\ 2. \qquad (16)$$

5.2.2. *Vanguard* 1.—For the orbit of Vanguard 1, which is very little affected by the atmosphere, the data given by Jacchia (1958) have been used. The rather more detailed information, extending over a longer time interval, which has subsequently been published (Kozai 1959a, O'Keefe & others 1959b), does not seem to be significantly different from the earlier data. The orbital elements given by Jacchia for 1958 June 18·9, are

$$i = 34{\cdot}253°, \qquad a/R = 1{\cdot}361527, \qquad e = 0{\cdot}19023,$$
$$\text{observed } \dot{\Omega} = -3{\cdot}0126(1\pm0{\cdot}0007)\ \text{deg/day}.$$

The luni-solar contribution to $\dot{\Omega}$ was shown by Kozai (1959b) to be $-0{\cdot}00041$ deg./day, and the effect of aerodynamic forces on $\dot{\Omega}$ is negligible. Subtracting the luni-solar effect from the observed $\dot{\Omega}$, taking a standard deviation of 0°·004 in i and 0·00001 in a/R, and substituting into equation (5), we have:

$$J_2-0{\cdot}682J_4-0{\cdot}017J_6 = 1083{\cdot}76\times10^{-6}(1\pm0{\cdot}0001)\text{—}\textit{Vanguard}\ 1. \qquad (17)$$

5.2.3. *Explorer* 7.—The Explorer 7 satellite, launched in October 1959, has a height of 550 km at perigee and about 1 100 km at apogee, and its orbit is so little affected by the atmosphere that average values over an interval of about six months can be used with negligible error in equation (5). In evaluating equation (5), three orbital parameters must be known with high accuracy: $\dot{\Omega}$, a and i. Values of Ω for Explorer 7 are issued regularly by Project Space Track, Bedford, Massachusetts, with their predictions. Since this satellite has been well observed, both optically and by radio, it is reasonable to hope that these values of Ω are accurate to about 0°·1 (standard deviation), and a scrutiny of the values shows that they are not inconsistent with this supposition. So, if an average value of $\dot{\Omega}$ over, say, 4 months is found by dividing the change in Ω, about 500°, by the time interval, we might hope that the error (s.d.) in $\dot{\Omega}$ would be as low as 1 in 3 000. Three values of $\dot{\Omega}$ were calculated in this way, end-points for Ω being chosen where ω was near 90° or 270°, thus making the average value of sin ω very near zero and virtually eliminating the effect of J_3 and J_5. The values of $-\dot{\Omega}$ obtained, corrected to 1960 January 11·5, were 4·1804, 4·1773 and 4·1785 deg/day, and a value of 4·1787 $(1\pm0{\cdot}0003)$ deg/day was taken. The orbital period T can be obtained correct to 0·005 min from the Space Track data, and hence, from the theoretical formula (King-Hele 1958) a can be found with a very small error. For 1960 January 11·5, a was taken as 7 198·32 $(1\pm0{\cdot}00003)$ km, and e, which does not need to be known so accurately, as 0·0370.

The third required orbital parameter of Explorer 7, the orbital inclination, was calculated from two observations by the author on 1960 February 15 and 16. In both observations, which were made from Farnham at a geocentric latitude of

51°·026, the zenith distance was less than 10°, the estimated standard deviation in direction 0°·1 (leading to s.d. of 0°·01 in i), and the satellite was within 0°·001 of its maximum latitude north, λ say. The values obtained for λ were 50°·274 and 50°·265, so that it seems legitimate to take $\lambda = 50°{\cdot}27 \pm 0°{\cdot}01$. Also, since λ differs from the smoothed inclination i by only

$$J_2 e\left(\frac{R}{p}\right)^2 \sin i \cos i(\cos\omega + \sin\omega) \simeq 0°{\cdot}001$$

(Merson 1960), i may be taken equal to λ. The value of 50°·27 for i is the same as that given by Space Track up to February 1960. Subsequently, however, values of i between 50°·29 and 50°·31 have been quoted by the U.S. prediction centres, and since there seems no reason for a secular increase in i, these subsequent values ought perhaps to be allowed to throw some semblance of a shadow before them. For the purposes of calculating the J_n, therefore, the value of i already quoted, $i = 50°{\cdot}27 \pm 0°{\cdot}01$, has been deliberately degraded in accuracy, and taken as $i = 50°{\cdot}28 \pm 0°{\cdot}02$, leading to an error factor of $(1 \pm 0{\cdot}0004)$ in $\cos i$.*

Using the values quoted above for $\dot{\Omega}$, a, e and i, and subtracting a luni-solar correction of $-0{\cdot}00015$ deg/day from $\dot{\Omega}$, equation (5) gives

$$J_2 + 0{\cdot}069J_4 - 0{\cdot}592J_6 = 1082{\cdot}14(1 \pm 0{\cdot}0005) \times 10^{-6} — \textit{Explorer} \text{ 7}. \qquad (18)$$

5.3. *Results*

Solving the simultaneous equations (16)–(18), we find

$$\left.\begin{aligned} J_2 &= (1082{\cdot}79 \pm 0{\cdot}15) \times 10^{-6} \\ J_4 &= (-1{\cdot}4 \pm 0{\cdot}2) \times 10^{-6} \\ J_6 &= (0{\cdot}9 \pm 0{\cdot}8) \times 10^{-6} \end{aligned}\right\}. \qquad (19)$$

It is noteworthy that the above values are consistent with the previous values, equation (14), which made use of a different satellite, Explorer 4.

6. Discussion

6.1. *The Earth's flattening*

If the Earth's sea-level surface were an exact spheroid, the flattening f' of the Earth, defined as the difference between the equatorial and polar diameters, divided by the equatorial diameter, would be given by the equation

$$f' = \tfrac{1}{2}(3J_2 + m)\left(1 + \frac{3}{4}J_2 + \frac{3}{28}m\right) + O(J_2^3) \qquad (20)$$

where $m = R^3w^2(1-f')/GM = 0{\cdot}0034498$ is the ratio of centrifugal to gravitational acceleration at the equator, multiplied by $(1-f')$, and w is the Earth's angular velocity. With $J_2 = (1082{\cdot}79 \pm 0{\cdot}15) \times 10^{-6}$, equation (20) gives $1/f' = 298{\cdot}24 \pm 0{\cdot}02$. Now that several higher-order harmonics have been evaluated,

* Subsequent determinations of the inclination of Explorer 7 (Y. Kozai, private communication) have yielded values lying between 50°·289 and 50°·320, and, if these are given due weight, the chosen value of i might be further amended to $50°{\cdot}29 \pm 0°{\cdot}02$. This would change the resulting values of J_2, J_4 and J_6 by about $\frac{1}{3}$ of their standard deviations, J_2 being increased, and $|J_4|$ and J_6 reduced.

the flattening is rather an outmoded concept; but it is still perhaps useful as a rough yardstick which brings a touch of reality into the mathematical abstractions of the potential.

6.2. *The accuracy of the results*

The errors caused by the neglect of the terms in J_7, J_8, ... are unknown: it cannot be proved that the cumulative effect of these terms is negligible. It is, however, intrinsically unlikely that any of the higher J_n is large: this would call for a bizarre variation of mass with latitude, which there is no reason to expect. Furthermore, the three sets of values of the even J's given in equations (14), (15) and (19) depend upon completely independent measurements—$\dot{\Omega}$ for Explorer 4, $\dot{\Omega}$ for Explorer 7, and $\dot{\omega}$ for Vanguard 1—and yet give consistent results, despite their neglect of the higher-order J_n. One might hope therefore that the errors incurred by neglecting the higher J_n do not greatly exceed the quoted standard deviations stemming from the observational errors. The J_n are a tempting target for the application of Occam's Razor: the way to progress here, it seems, is to take the simplest hypothesis consistent with the facts, though in the foreknowledge that it may later be modified.

The question-mark created by the neglect of the higher J_n will remain even when both theory and observation have been greatly improved, and many more of the J_n have been evaluated. Indeed, the higher the accuracy of the determinations, the more boldly the question-mark will stand out. The only way to avoid it would be to invent a new mathematical form, quite different from equation (2), for the expansion of the potential. It may be that before such a radical change becomes necessary the series of values obtained for the early J_n will show an undeniable tendency to diminish. The present set of values,

$$\left.\begin{aligned} J_2 &= (1\,082{\cdot}79 \pm 0{\cdot}15) \times 10^{-6} \\ J_3 &= (-2{\cdot}4 \pm 0{\cdot}3) \times 10^{-6} \\ J_4 &= (-1{\cdot}4 \pm 0{\cdot}2) \times 10^{-6} \\ J_5 &= (-0{\cdot}1 \pm 0{\cdot}1) \times 10^{-6} \\ J_6 &= (0{\cdot}9 \pm 0{\cdot}8) \times 10^{-6} \end{aligned}\right\} \qquad (21)$$

could be invoked to support either of two conflicting hypotheses: (i) that the J_n quickly become smaller as n increases; or (ii) that J_n is likely to be of order $0{\cdot}2 \times 10^{-6}$, however large n becomes. Sir Harold Jeffreys (private communication) has however pointed out that hypothesis (ii) is untenable since the sum of the squares of the terms on the r.h.s. of equation (2) would be infinite, and he argues from considerations of continuity that the J_n should decrease at least as fast as $1/n$.

It seems fair to say, in conclusion, that the standard deviations in equations (21) are as realistic as such estimates ever can be in our inevitable ignorance about future discoveries which would influence our views. Even if the errors are rather larger than quoted, the evaluation of so many of the J_n in the last two years has represented something of a quantum jump in the process of advance from the first determinations of J_2, by Newton, Maupertuis and La Condamine, to the day when the values of the J_n will have to be revised after every major earthquake.

Acknowledgments

I wish to thank Sir Harold Jeffreys for several valuable comments on the draft version of this paper; Mr R. H. Merson for making available the results of his orbital theory prior to publication and for helpful discussions; and Mrs D. M. C. Walker for her assistance with the calculations and for some searching criticisms.

Royal Aircraft Establishment,
Farnborough, Hants:
1960 *August.*

References

Bomford, G., 1960. Report of artificial satellite study group. *Geophys. J.* To be published.

Brouwer, D., 1959. Solution of the problem of artificial satellite theory without drag. *Astronom. J.*, **64**, 378–397.

Buchar, E., 1958. Motion of the nodal line of the second Russian earth satellite, and flattening of the Earth. *Nature, Lond.*, **182**, 198–199.

Buchar, E., 1960. Determination of some parameters of the gravity field of the Earth from the rotation of the nodal line of artificial satellites. (Paper circulated at IUGG Assembly, Helsinki, August 1960.)

Cohen, C. J. & Anderle, R. J., 1960. Verification of the Earth's "pear-shape" gravitational potential. *Science*, **131**, 807–808.

Cook, A. H., 1959. Developments in dynamical geodesy. *Geophys. J.*, **2**, 222–240.

Cook, A. H., 1960. Resonant orbits of artificial satellites. *Space Research*, Proc. 1st Int. Space Sci. Symposium, p. 476. North Holland Publ. Co.

Cook, G. E., 1960. The effect of an oblate, rotating atmosphere on the orientation of a satellite orbit. Ministry of Aviation Report (to be published).

Cornford, E. C., 1958. A comparison of orbital theory with observations made in the United Kingdom on the Russian satellites. Paper presented at 5th C.S.A.G.I. Assembly, Moscow, 1958.

Groves, G. V., 1960. Motion of a satellite in the Earth's gravitational field. *Proc. Roy. Soc. A*, **254**, 48–65.

Harvard, 1958. Oblateness of the Earth by artificial satellites. Harvard Announcement Card 1408, 1958 June 24.

Jacchia, L. G., 1958. The Earth's gravitational potential as derived from satellites 1957 β 1 and 1958 β 2. Smithsonian Astrophys. Obs. Special Report No. 19.

Jeffreys, H., 1952. *The Earth.* Cambridge U.P. 3rd edition.

King-Hele, D. G., 1958. The effect of the Earth's oblateness on the orbit of a near satellite. *Proc. Roy. Soc. A*, **247**, 49–72.

King-Hele, D. G., 1959a. Analysis of the orbits of the Russian satellites. Paper presented at Royal Society conference, 1958 November 12. *Proc. Roy. Soc. A*, **253**, 529–538.

King-Hele, D. G., 1959b. Determination of air density and the Earth's gravitational field from the orbits of artificial satellites. *Proc. 10th Int. Ast. Cong.*, London, 1959. Springer Verlag, Vienna, pp. 1–20.

King-Hele, D. G. & Merson, R. H., 1958. Satellite orbits in theory and practice. *J. Brit. Interplan. Soc.*, **16**, 446–471.

King-Hele, D. G. & Merson, R. H., 1959. A new value for the Earth's flattening, derived from measurements of satellite orbits. *Nature, Lond.*, **183**, 881–882.

Kozai, Y., 1959a. The Earth's gravitational potential derived from the motion of satellite 1958 β 2. Smithsonian Astrophys. Obs. Special Report No. 22, pp. 1–5.

Kozai, Y., 1959b. On the effects of the Sun and the Moon upon the motion of a close Earth satellite. Smithsonian Astrophys. Observatory Special Report No. 22, pp. 7–10.

Kozai, Y., 1959c. The motion of a close Earth satellite. *Astronom. J.*, **64**, 367–377.

Kozai, Y., 1960. The gravitational field of the Earth derived from the motions of three satellites. *Astronom. J.* (to be published).

Krause, H. G. L., 1956. Die Säkularen und periodischen Störungen der Bahn eines Künstlichen Erdsatelliten. *Procceedings of the 7th Int. Astronautical Cong., Rome*, 1956, pp. 523–585.

Lecar, M., Sorenson, J. & Eckels, A., 1959. A determination of the coefficient J of the second harmonic in the earth's gravitational potential from the orbit of satellite 1958β2. *J. Geophys. Res.*, **64**, 209–216.

Merson, R. H., 1960. The motion of a satellite in an axisymmetric gravitational field. *Geophys. J.*, **4**, 17.

Merson, R. H. & King-Hele, D. G., 1958. Use of artificial satellites to explore the Earth's gravitational field: results from Sputnik 2. *Nature, Lond.*, **182**, 640–641.

Message, P. J., 1960. On Mr King-Hele's theory of the effect of the earth's oblateness on the orbit of a close satellite. *Mon. Not. R. Astr. Soc.*, **121**, 1–4.

Musen, P., 1960. The influence of solar radiation pressure on the motion of an artificial satellite. *J. Geophys. Res.*, **65**, 1391–1396.

O'Keefe, J., Eckels, A. & Squires, R. K., 1959a. Vanguard measurements give pear-shaped component of Earth's figure. *Science*, **129**, 565–566.

O'Keefe, J. A., Eckels, A. & Squires, R. K., 1959b. The gravitational field of the Earth. *Astronom. J.*, **64**, 245–253.

Roberson, R. E., 1957. Orbital behaviour of Earth satellites. *J. Franklin Inst.*, **264**, 181–201 and 269–285.

Zhongolovitch, I. D., 1960a. Formulae relating to the motion of a particle in the gravitational field of a level ellipsoid of revolution. *Bulletin of the U.S.S.R. Institute of Theoretical Astronomy*, **7**, 521–536.

Zhongolovitch, I. D., 1960b. An attempt to determine certain parameters of he Earth's gravitational field from the results of observations of satellites 1957β2 1958δ1 and 1958δ2. Bulletin of Stations for Optical Observation of Artificial Earth Satellites, No. 2, 1960, pp. 1–24.

The Motion of a Satellite in an Axi-symmetric Gravitational Field

R. H. Merson

"Now fills thy sleep with perturbations."—*W. Shakespeare, Richard III*, V, 3.

Summary

The equations for the variation of the osculating elements of a satellite are integrated to yield the complete perturbations of the first order due to the second harmonic, together with the secular perturbations of the second order due to the second harmonic and of the first order due to the third to sixth harmonics. A set of smoothed elements is then derived, in which the perturbations of the even harmonics have no singularities, the semi-major axis and eccentricity have no variation due to the second harmonic and the other elements have the smallest possible amplitudes of oscillation. The formulae presented will be extremely useful in the reduction of earth-satellite observations and geopotential studies based on these.

1. Introduction

In the widening fields of research opened up by the launching of Earth satellites there are several investigations which depend for their fulfilment on an accurate knowledge of the motion of a satellite in the vicinity of the Earth. The determination of upper atmosphere densities (King-Hele & Walker 1960) and the gravitational field of the Earth (Merson & King-Hele 1958, 1960; Cook 1958) are two topics which have already received considerable attention.

With the advent of observations accurate to a few seconds of arc, such as those provided by Baker-Nunn cameras, the necessity has arisen for more accurate theories of satellite motion than have previously been used. For example, it is no longer satisfactory to assume that the inclination of the orbital plane remains constant during a few revolutions, for the maximum northern latitude reached can differ from the maximum southern latitude reached by more than 20 seconds of arc on the same revolution. This fact cannot be obscured by altering the definitions of the elements in any satisfactory manner.

The simplest analytical approach to the problem of satellite motion lies in the use of osculating elements, which are easily understood and have a long history of use in astronomy. However, they are not suitable for use as working elements in the case of a close Earth-satellite for two main reasons: the amplitudes of the oscillations in the elements are relatively large (Brenner 1959), and the perturbation in the argument of perigee has a factor $(\text{eccentricity})^{-1}$ which becomes unmanageable in the case of near-circular orbits.

The procedure adopted in this paper is to derive first the perturbations of the osculating elements, and then to consider transformations of the elements which reduce the short-period perturbations to the smallest possible amplitudes. It is shown, in fact, that there exists a set of elements of a particularly simple form in which semi-major axis a and eccentricity e remain constant when the second harmonic alone is considered.

List of Symbols

a,	semi-major axis	S,	$= \sin u$
C,	$= \cos u$	t,	time, measured from the ascending node
e,	eccentricity	T_Ω,	nodal period
E,	eccentric anomaly	u,	argument of the latitude
$\mathbf{f}$,	acceleration vector of perturbing force	U,	gravitational potential
f_1, f_2, f_3	components of $\mathbf{f}$	v,	true anomaly
f,	$= \sin^2 i$	V,	satellite speed
F,	perturbing potential	z,	variable representing radial perturbation
H,	angular momentum		
i,	orbital inclination	ζ,	arbitary element
J_2, J_3, etc.,	coefficients in the potential expansion	θ,	co-latitude
		ϕ,	auxiliary element
M,	mean anomaly	λ,	right ascension
n,	mean motion	μ,	gravitational constant
p,	latus rectum of orbit	σ,	mean anomaly at the node
$\mathbf{r}$,	radius vector of satellite	σ^*,	modified mean anomaly
r,	magnitude of $\mathbf{r}$	ω,	argument of perigee
R,	mean equatorial radius of Earth	Ω,	right ascension of the node

2. Osculating elements

In the absence of a perturbing force, the motion of a negligibly small particle in a field with potential $U = \mu/r$ is an ellipse. Here, r is the distance from the centre of attraction and μ is a constant.

When a perturbing force is acting, the osculating elliptic orbit at any given time is defined as the orbit which the particle (see Figure 1) would have if the perturbing

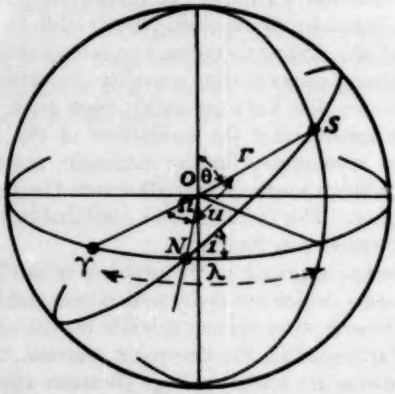


Fig. 1.—The projection of the osculating orbit on the celestial sphere.
γ = first point of Aries, N = ascending node, S = satellite.

force suddenly ceased to exist. The osculating orbit is specified by six independent parameters or elements which will be taken to be the semi-major axis a, the eccentricity e, the inclination i to the equator, the right ascension Ω of the ascending node, the argument of perigee ω, and σ^*, the modified mean anomaly, which is related to σ, the mean anomaly at time $t = 0$ by the relation

$$\sigma^* = \sigma + nt - \int_0^t n\,dt, \tag{1}$$

where n is the mean motion, defined by the equation

$$n^2a^3 = \mu, \tag{2}$$

and is not in general constant.

At any given time the polar coordinates r, θ, λ of the particle are related to the orbital elements by the well-known formulae

$$r = a(1 - e\cos E) = p/(1 + e\cos v), \tag{3}$$

$$\sin\theta\cos(\lambda - \Omega) = \cos u, \tag{4}$$

$$\sin\theta\sin(\lambda - \Omega) = \cos i\sin u, \tag{5}$$

$$\cos\theta = \sin i\sin u, \tag{6}$$

where θ is the co-latitude and λ the right ascension of the satellite; u is the argument of the latitude and v the true anomaly, and they are related by

$$u = v + \omega; \tag{7}$$

E is the eccentric anomaly, which is related to the true anomaly v, the mean anomaly M and time t by the equations

$$\cos E = (e + \cos v)/(1 + e\cos v), \tag{8}$$

$$\sin E = (1 - e^2)^{\frac{1}{2}}\sin v/(1 + e\cos v), \tag{9}$$

$$M = E - e\sin E = nt + \sigma = \int_0^t n\,dt + \sigma^*; \tag{10}$$

and finally

$$p = a(1 - e^2) \tag{11}$$

is the latus-rectum of the orbit.

3. The time rates of change of the elements

If the acceleration vector due to the perturbing force is denoted by $\mathbf{f}$, then the motion of the particle is governed by the equation

$$\frac{d^2\mathbf{r}}{dt^2} = -\frac{\mu\mathbf{r}}{r^3} + \mathbf{f},$$

where $\mathbf{r}$ is the position vector of the particle relative to the centre of force.

Let f_1 be the component of $\mathbf{f}$ along $\mathbf{r}$, f_2 the component perpendicular to $\mathbf{r}$ in the osculating plane and in the direction of increasing anomaly, and f_3 the component perpendicular to the osculating plane in the direction of the angular momentum vector. Then (Smart 1953, para. 14.04 slightly modified)

$$\dot{a} = 2(a^2/H)(f_1 e \sin v + f_2 p/r), \tag{12}$$

$$\dot{e} = (H/\mu)[f_1 \sin v + f_2 \cos v(1 + r/p) + f_2 e r/p] \tag{13}$$

$$di/dt = r f_3 \cos u/H \tag{14}$$

$$\dot{\Omega} = r f_3 \sin u/(H \sin i) \tag{15}$$

$$\dot{\omega} = (H/\mu e)[-f_1 \cos v + f_2 \sin v(1 + r/p) - f_3 e \sin u \cot i(r/p)] \tag{16}$$

$$\dot{\sigma}^* = (p/na^2 e)[f_1(\cos v - 2er/p) - f_2 \sin v(1 + r/p)], \tag{17}$$

where $H = (\mu p)^{\frac{1}{2}}$ is the angular momentum of the particle about the centre of force.

Since it is more convenient to use u as independent variable, the derivative du/dt is required.

Tisserand (1889, p. 462) gives

$$r^2(\dot{u} + \dot{\Omega} \cos i) = H \tag{18}$$

which, combined with (15), gives

$$\dot{u} = (H/r^2) - r f_3 \sin u \cot i/H. \tag{19}$$

Also, combining (15), (16) and (17)

$$(1 - e^2)^{\frac{1}{2}} \dot{\sigma}^* + \dot{\omega} + \dot{\Omega} \cos i = -(1 - e^2)^{\frac{1}{2}} \dot{\phi}, \tag{20}$$

where

$$\dot{\phi} = 2nar f_1/\mu, \tag{21}$$

ϕ being introduced as a useful auxiliary element.

A detailed proof *ab initio* of all the formulae given above is to be given in a fuller account of the present work, to be published shortly.

4. The perturbations due to an axi-symmetric gravitational field

If the only forces acting on a satellite (hereinafter used instead of "particle") are those due to a gravitational field with axial symmetry, and if the origin of coordinates is suitably chosen, the potential can be expressed as a series of spherical harmonics

$$U = (\mu/r)[1 - \sum_{n=2}^{\infty} J_n (R/r)^n P_n(\cos \theta)] \tag{22}$$

where J_n, R are constants, θ is the angle between the radius vector $\mathbf{r}$ and the axis of symmetry and P_n is the Legendre polynomial of order n.

The perturbing potential, denoted by F, is thus

$$F = U - \mu/r = -(\mu/r) \sum_{n=2}^{\infty} J_n (R/r)^n P_n(\cos \theta). \tag{23}$$

The component of the perturbing acceleration **f** in any direction is the rate of change of F in that direction. With the aid of simple spherical trigonometry based on Figure 1 it follows that

$$\left.\begin{aligned} f_1 &= \partial F/\partial r \\ f_2 &= -\left(\frac{\cos u \sin i}{r \sin\theta}\right)\frac{\partial F}{\partial \theta} \\ f_3 &= -\left(\frac{\cos i}{r\sin\theta}\right)\frac{\partial F}{\partial \theta}. \end{aligned}\right\} \tag{24}$$

Substituting the explicit formulae for P_n up to $n = 6$, performing the differentiations, then setting $\cos\theta = \sin i \sin u$, and using the abbreviations $S \equiv \sin u$, $f \equiv \sin^2 i$, we find

$$f_1 = (\mu/r^2)\left[\frac{3}{2}J_2(R/r)^2(3fS^2-1)+2J_3(R/r)^3\sin i(5fS^3-3S)+\right.$$
$$+\frac{3}{8}J_4(R/r)^4(35f^2S^4-30fS^2+3)+\frac{3}{4}J_5(R/r)^5\sin i(63f^2S^5-70fS^3+15S)+$$
$$\left.+\frac{7}{16}J_6(R/r)^6(231f^3S^6-315f^2S^4+105fS^2-5)\right], \tag{25}$$

$$f_2 = -(\mu/r^2)\sin i\cos u\left[3J_2(R/r^2)\sin i\sin u+\frac{3}{2}J_3(R/r)^3(5fS^2-1)+\right.$$
$$+\frac{5}{2}J_4(R/r)^4\sin i(7fS^2-3S)+\frac{15}{8}J_5(R/r)^5(21f^2S^4-14fS^2+1)+$$
$$\left.+\frac{21}{8}J_6(R/r)^6\sin i(33f^2S^5-30fS^3+5)\right], \tag{26}$$

$$f_3 = -(\mu/r^2)\cos i\times \text{ [same as } f_2\text{]}. \tag{27}$$

From (23), the explicit expansion of the perturbing potential F up to $n = 6$ is

$$F = -(\mu/r)\left[\tfrac{1}{2}J_2(R/r)^2(3fS^2-1)+\tfrac{1}{2}J_3(R/r)^3\sin i(5fS^3-3S)+\right.$$
$$+\tfrac{1}{8}J_4(R/r)^4(35f^2S^4-30fS^2+3)+\tfrac{1}{8}J_5(R/r)^5\sin i(63f^2S^5-70fS^3+15S)+$$
$$\left.+\tfrac{1}{16}J_6(R/r)^6(231f^3S^6-315f^2S^4+105fS^2-5)\right]. \tag{28}$$

5. Two exact integrals

Two exact integrals can be found for a dynamical system possessing a potential function, the integrals of energy and of area.

Now the kinetic energy of the satellite per unit mass is $\frac{1}{2}V^2$, where $V = |\dot{\mathbf{r}}|$, and the potential energy is $-U$, so that the total energy is $\frac{1}{2}V^2 - U$, which is constant.

But (Smart 1953, para. 2.12)

$$V^2 = \mu\left(\frac{2}{r} - \frac{1}{a}\right) \tag{29}$$

for an unperturbed orbit; and since the osculating orbit is a function of the two vectors $\mathbf{r}, \dot{\mathbf{r}}$ only, equation (29) may be regarded as defining a in the perturbed orbit.

Thus, using (29) and (23), the total energy is

$$\left(\frac{\mu}{r} - \frac{\mu}{2a}\right) - \left(F - \frac{\mu}{r}\right),$$

so that

$$\frac{\mu}{2a} + F = \text{constant}. \tag{30}$$

This equation enables the perturbation in a to be written down without explicitly integrating equation (12).

To determine the second integral it is noted first that, from (11), (12) and (13),

$$\dot{p} = 2Hrf_2/\mu.$$

Also, from (26) and (27),

$$f_2 = f_3 \tan i \cos u.$$

From these two equations and (14) it follows that

$$\frac{d}{dt}(p\cos^2 i) = \cos^2 i \cdot \frac{dp}{dt} - 2p\cos i \sin i \frac{di}{dt} = 0$$

and hence

$$p\cos^2 i = \text{constant}. \tag{31}$$

It is not immediately obvious that this is related to an integral of area, but a simple computation will show it.

Differentiating (4) and (6) and eliminating $\dot{\theta}$ gives

$$\dot{\lambda} = \dot{\Omega} + \dot{u}\cos i \operatorname{cosec}^2\theta - \frac{di}{dt}\sin i \sin u \cos u \operatorname{cosec}^2\theta, \tag{32}$$

and substituting for $\dot{\Omega}$, $\dot{u}$, di/dt from (15), (19) and (14) then gives

$$\dot{\lambda} = H\cos i / r^2 \sin^2\theta,$$

or

$$\dot{\lambda} r^2 \sin^2\theta = (\mu p \cos^2 i)^{\frac{1}{2}}. \tag{33}$$

Now $r\sin\theta$ is the projection of the radius vector on the equatorial plane, and λ is the angle turned through by this projection: so that $\dot{\lambda} r^2 \sin^2\theta$ is the rate of change of area swept out by the projection, and (33), (31) show that this is constant.

6. First order J_2 terms

Since, for the Earth, J_2 is about 10^{-3} and $J_3, J_4 \ldots$ are of order 10^{-6} or less, the major part of the perturbation is contained in the first-order J_2 terms. In the case of the semi-major axis no integration is required, for from (30)

$$\frac{\mu}{2a}+F=\frac{\mu}{2a_0}+F_0,$$

where the suffix 0 is used to denote conditions at the ascending node ($t=0$, $u=0$).

The symbol δ will be used to denote departures from nodal conditions, so that

$$\delta a = a-a_0 = 2aa_0(F-F_0)/\mu, \tag{34}$$

or, substituting the J_2 term from (28),

$$\delta a = J_2R^2aa_0[r^{-3}(1-3fS^2)-r_0^{-3}].$$

Finally, substituting for r from (9), noting that $v=-\omega$ at $u=0$ and also that $a_0=a+O(J_2)$, we have

$$\delta a = J_2R^2a^2p^{-3}[(1+e\cos v)^3(1-3fS^2)-(1+e\cos\omega)^3]+O(J_2^2). \tag{35}$$

The procedure for the remaining elements is (i) to substitute the J_2 terms of f_1, f_2 and f_3 in equations (13)–(17); (ii) to change the independent variable to u using

$$\frac{du}{dt}=\frac{H}{r^2}+O(J_2);$$

(iii) to substitute for $r=p/(1+e\cos v)$, after which it is seen that the right-hand sides are functions of H, p, n, a, e, i, u and $v=u-\omega$. Since H, p, n, a, e, i, ω are constant $+O(J_2)$, the rates of change of the elements with respect to u are, to the first order in J_2, functions of u only and are readily integrable.

After considerable manipulation the results below have been obtained, with $f\equiv\sin^2 i$, $S\equiv\sin u$ and $C\equiv\cos u$ for brevity. There are many different ways in which these formulae can be presented; for example, $\cos\omega$ can be replaced by $C\cos v+S\sin v$ and $\cos^2 v$ by $\frac{1}{2}+\frac{1}{2}\cos 2v$. As the main use of the formulae is for further analytical work, both in determining the secular (J_2^2) terms and later in the transformation of the elements, the presentation is given with this work in mind. The curly brackets are used to separate odd-periodic from even-periodic terms, that is terms in which the sum of the powers of S, C, $\sin v$ and $\cos v$ are either odd or even.

$$\begin{aligned}\delta e = {} & 3J_2(R/p)^2\times\\
&\times\Big[\Big\{\Big(-\frac{1}{2}+\frac{1}{3}f\Big)\cos\omega-\frac{1}{4}e(1+\cos 2\omega)+\Big(-\frac{1}{8}-\frac{1}{3}f\Big)e^2\cos\omega-\frac{1}{24}e^2\cos 3\omega-\\
&-\frac{1}{2}feS^2-\frac{3}{2}feS^2\cos^2 v+\frac{1}{2}e\cos^2 v\Big\}+\\
&+\Big\{-\frac{1}{3}f(1-e^2)C\cos\omega+\frac{1}{2}\cos v-\Big(\frac{7}{6}+\frac{1}{3}e^2\Big)fS^2\cos v-\\
&-\frac{1}{2}fe^2S^2\cos^3 v+\frac{1}{6}e^2\cos^3 v\Big\}\Big]+O(J_2^2),\end{aligned} \tag{36}$$

$$\delta i = -3J_2(R/p)^2 \sin i \cos i \left[\left\{\tfrac{1}{2}S^2 + \tfrac{1}{3}e\cos\omega\right\} + \left\{\tfrac{1}{3}S^2 e\cos v - \tfrac{1}{3}Ce\cos\omega\right\}\right] + O(J_2^2), \quad (37)$$

$$\delta\Omega = -3J_2(R/p)^2 \cos i \times$$

$$\times \left[\tfrac{1}{2}u + \left\{\tfrac{2}{3}e\sin\omega - \tfrac{1}{2}SC\right\} + \left\{\tfrac{1}{3}S^2 e\sin v - \tfrac{2}{3}Ce\sin\omega\right\}\right] + O(J_2^2), \quad (38)$$

$$\delta\omega = 3J_2(R/p)^2 \times$$

$$\times \left[\left(1 - \tfrac{5}{4}f\right)u + \left\{\left(\tfrac{1}{2} - \tfrac{1}{3}f\right)e^{-1}\sin\omega + \left(\tfrac{25}{24} - \tfrac{4}{3}f\right)e\sin\omega + \tfrac{1}{24}e\sin 3\omega + \right.\right.$$

$$+ \left(\tfrac{5}{4}f - \sin^2\omega\right)SC - \tfrac{3}{2}fS^2\sin v\cos v + \tfrac{1}{2}S^2\sin 2\omega\Big\} +$$

$$+ \left\{\left(\tfrac{1}{3}fe^{-1} - \tfrac{2}{3}e + \tfrac{4}{3}fe\right)C\sin\omega + \left(\tfrac{1}{3}e - \tfrac{7}{6}fe^{-1} - \tfrac{7}{6}fe\right)S^2\sin v + \right.$$

$$\left.\left. + \tfrac{1}{2}feS^2\sin^3 v + \tfrac{1}{2}(e^{-1} + e)\sin v - \tfrac{1}{6}e\sin^3 v\right\}\right] + O(J_2^2), \quad (39)$$

$$\delta\sigma^* = 3J_2(R/p)^2(1-e^2)^{\frac{1}{2}} \times$$

$$\times \left[\left(\tfrac{1}{2} - \tfrac{3}{4}f\right)u - \left\{\left(\tfrac{1}{2} - \tfrac{1}{3}f\right)e^{-1}\sin\omega + \left(-\tfrac{5}{8} + \tfrac{4}{3}f\right)e\sin\omega + \tfrac{1}{24}e\sin 3\omega + \right.\right.$$

$$+ \left(\tfrac{1}{2}\cos 2\omega - \tfrac{3}{4}f\right)SC - \tfrac{3}{2}fS^2\sin v\cos v + \tfrac{1}{2}S^2\sin 2\omega\Big\} -$$

$$- \left\{\left(\tfrac{1}{3}fe^{-1} - \tfrac{4}{3}fe\right)C\sin\omega + \left(\tfrac{1}{6}fe - \tfrac{7}{6}fe^{-1}\right)S^2\sin v + \right.$$

$$\left.\left. + \tfrac{1}{2}feS^2\sin^3 v + \tfrac{1}{2}(e^{-1} - e)\sin v - \tfrac{1}{6}e\sin^3 v\right\}\right] + O(J_2^2). \quad (40)$$

Besides these, the perturbations in p and the auxiliary element ϕ will be required. Since $p\cos^2 i$ = constant (33), the perturbation in p is easily found: for

$$\delta p = 2p\tan i\,\delta i + O(J_2^2)$$

$$= -3J_2(R/p)^2 f\left[\left\{S^2 + \tfrac{2}{3}e\cos\omega\right\} + \left\{\tfrac{2}{3}S^2 e\cos v - \tfrac{2}{3}Ce\cos\omega\right\}\right] + O(J_2^2). \quad (41)$$

$\delta\phi$ is obtained from equation (21) by applying the same procedure as that applied to the other elements:

$$\delta\phi = 3J_2(R/p)^2(1-e^2)^{\frac{1}{2}} \times$$

$$\times \left[\left(\tfrac{3}{2}f - 1\right)u + \left\{2fe\sin\omega - \tfrac{3}{2}fSC - e\sin\omega\right\} + \right.$$

$$\left. + \left\{feS^2\sin v - 2feC\sin\omega - e\sin v\right\}\right] + O(J_2^2). \quad 42)$$

Two checks have been applied which between them, cover all the above formulae. First, from (11),

$$\delta p = (1-e^2)\delta a - 2ea\delta e + O(J_2^2) \tag{43}$$

and second, from (20),

$$(1-e^2)^{-\frac{1}{2}}(\delta\sigma^* + \delta\phi) = -(\delta\omega + \delta\Omega\cos i) + O(J_2^2). \tag{44}$$

7. The maximum latitude

One feature of the above formulae which has not been commented upon before is the difference between the maximum northerly and southerly latitudes reached by the satellite. At these points $\dot{\theta} = 0$.

Differentiating equation (6) with respect to time and substituting for $\dot{u}$ from (19) and di/dt from (14),

$$-\sin\theta \,.\, \dot{\theta} = H\sin i\cos u/r^2.$$

Thus, apart perhaps from degenerate cases, $\dot{\theta} = 0$ when $\cos u = 0$, i.e. at $u = \pi/2$ for maximum north-going and $u = 3\pi/2$ for maximum south-going excursions.

Substitution in (37) then gives

$$(i-i_0)_{u=\pi/2} = -3J_2(R/p)^2\sin i\cos i\left[\frac{1}{2}+\frac{1}{3}e\cos\omega+\frac{1}{3}e\sin\omega\right]+O(J_2^2)$$

and

$$(i-i_0)_{u=3\pi/2} = -3J_2(R/p)^2\sin i\cos i\left[\frac{1}{2}+\frac{1}{3}e\cos\omega-\frac{1}{3}e\sin\omega\right]+O(J_2^2).$$

The difference between these is

$$i_{u=3\pi/2} - i_{u=\pi/2} = \frac{1}{2}J_2(R/p)^2\sin 2i \,.\, e\sin\omega + O(J_2^2), \tag{45}$$

which for $i = 45°$, $e = 0{\cdot}1$, $\omega = 90°$ is greater than 20 seconds of arc for near Earth-satellites.

This difference must therefore be taken into account if the most accurate observations are to be used efficiently.

8. Secular and long-period terms of the first order (J_2 to J_6)

To determine the secular and long-period terms of the first order, the expressions (25)–(27) for f_1, f_2, f_3 are substituted in equations (13)–(17) which are then multiplied by $dt/du = r^2/H + O(J_2)$. If ζ stands for any of the elements, then $d\zeta/du$ is seen to be expressible as the sum of a number of terms of the general form

$$\Phi(i, e, \omega)\sin^l u\cos^m u\sin^q v\cos^s v \tag{46}$$

where n, a, have been eliminated through equations (2) and (11), H has been set equal to $(\mu p)^{\frac{1}{2}}$ and p eliminated through equation (33). l, m, q, s are positive integers or zero.

The first-order perturbation is then obtained by integrating $d\zeta/du$ with i, e, ω kept constant.

The secular and long-period changes in the elements are obtained by integrating from $u = 0$ to $u = 2\pi$, i.e. from one ascending node to the next. The result is denoted by $\Delta\zeta$, so that

$$\Delta\zeta = \zeta_{2\pi} - \zeta_0 = \int_0^{2\pi} \frac{d\zeta}{du} du. \tag{47}$$

A full list of the basic integrals required is to be published elsewhere. It will merely be noted here that the expression (46) contributes nothing to $\Delta\zeta$ if $l+m+q+s$ is odd so that, in forming $d\zeta/du$, odd-periodic terms can be omitted.

The results are tabulated at the end of the next section.

9. Secular and long-period terms of the second order (J_2^2)

The only second-order perturbations of any consequence are those of the type J_2^2, for $J_2^2 \approx 10^{-6}$ and is of the same order as $J_3, J_4 \ldots$.

To determine the J_2^2 perturbations, the J_2 terms in f_1, f_2, f_3 are substituted in equations (13)–(17) which are then multiplied by

$$\frac{dt}{du} = \frac{r^2}{H}\left(1 - \frac{3J_2R^2}{pr}\cos^2 i \sin^2 u\right) + O(J_2^2). \tag{48}$$

This time the elements i, e, ω in the expression for $d\zeta/du$ are not regarded as constants, but their first-order variations, as given by (36), (37) and (39), are used.

More explicitly, let

$$X_\zeta = \left(\frac{r^2}{H}\frac{d\zeta}{dt}\right)_{i=i_0, e=e_0, \omega=\omega_0, J_3=J_4=\ldots=0}, \tag{49}$$

and let

$$Y_\zeta = (-3J_2(R^2/pr)\cos^2 i \sin^2 u\, X_\zeta)_{i=i_0, e=e_0, \omega=\omega_0} \tag{50}$$

so that X_ζ, Y_ζ are functions of i_0, e_0, ω_0 and u.

It follows that

$$\frac{d\zeta}{du} = X_\zeta + \frac{\partial X_\zeta}{\partial i_0}.\delta i + \frac{\partial X_\zeta}{\partial e_0}.\delta e + \frac{\partial X_\zeta}{\partial \omega_0}.\delta\omega + Y_\zeta + O(J_2^3, J_3 \ldots), \tag{51}$$

where $X_\zeta = O(J_2)$ and the other terms are $O(J_2^2)$, and so

$$\Delta\zeta = \text{1st order terms} + \int_0^{2\pi}\left[\delta i\frac{\partial X_\zeta}{\partial i_0} + \delta e\frac{\partial X_\zeta}{\partial e_0} + \delta\omega\frac{\partial X_\zeta}{\partial \omega_0}\right]du + \int_0^{2\pi} Y_\zeta\, du +$$

$$+ O(J_2^3, J_3^2, J_4^2, \ldots, J_2J_3, \ldots). \tag{52}$$

These integrals are extremely laborious to calculate, especially for $\zeta = \omega$ and σ^*, and the necessity for good checks is paramount. Fortunately such a check is

available, at the expense of doubling the amount of calculation, for if ν is any one of i, e, ω, integration by parts gives

$$\int_0^{2\pi} \delta\nu \cdot \frac{\partial X_\zeta}{\partial \nu_0}\,du = (\delta\nu)_{2\pi} \cdot \int_0^{2\pi} \frac{\partial X_\zeta}{\partial \nu_0}\,du - \int_0^{2\pi} \frac{d}{du}(\delta\nu)\left(\int_0^u \frac{\partial X_\zeta}{\partial \nu_0}\,du\right)du$$

$$= \Delta\nu \cdot \frac{\partial}{\partial \nu_0}\int_0^{2\pi} X_\zeta\,du - \int_0^{2\pi} X_\nu\left(\frac{\partial}{\partial \nu_0}\int_0^u X_\zeta\,du\right)du$$

$$= \Delta\nu \cdot \frac{\partial}{\partial \nu_0}\int_0^{2\pi}\left(\frac{d\zeta}{du}+O(J_2^2)\right)du - \int_0^{2\pi} X_\nu\left\{\frac{\partial}{\partial \nu_0}\int_0^u\left(\frac{d\zeta}{du}+O(J_2^2)\right)\right\}du$$

$$= \Delta\nu \cdot \frac{\partial}{\partial \nu_0}(\Delta\zeta) - \int_0^{2\pi} X_\nu \frac{\partial}{\partial \nu_0}(\delta\zeta)\,du + O(J_2^3).$$

Also, since $\Delta e = O(J_2^2)$, $\Delta i = O(J_2^2)$, the first term on the right is $O(J_2^3)$ when $\nu = e$ or i. Thus

$$\left.\begin{aligned}
\int_0^{2\pi} \delta i \cdot \frac{\partial X_\zeta}{\partial i_0}\,du &= -\int_0^{2\pi} X_i \frac{\partial}{\partial i_0}(\delta\zeta)\,du + O(J_2^3)\\
\int_0^{2\pi} \delta e \cdot \frac{\partial X_\zeta}{\partial e_0}\,du &= -\int_0^{2\pi} X_e \frac{\partial}{\partial e_0}(\delta\zeta)\,du + O(J_2^3)\\
\int_0^{2\pi} \delta\omega \cdot \frac{\partial X_\zeta}{\partial \omega_0}\,du &= \Delta\omega \cdot \frac{\partial}{\partial \omega_0}(\Delta\zeta) - \int_0^{2\pi} X_\omega \frac{\partial}{\partial \omega_0}(\delta\zeta)\,du + O(J_2^3).
\end{aligned}\right\} \quad (53)$$

These provide a very powerful check, because it is found that the basic integrals (of expressions (46)) enter in different proportions, and so they themselves are also checked.

The main results are tabulated below.

No integration is required to determine Δa, for from (34)

$$\begin{aligned}
\Delta a &= 2a_{2\pi}a_0\mu^{-1}(F_{2\pi}-F_0)\\
&= J_2R^2a_0^2(r_{2\pi}^{-3}-r_0^{-3})+O(J_2^3,J_3^2,\ldots),
\end{aligned}$$

and since Δe, Δp have no first-order secular J_2 variation,

$$\begin{aligned}
\Delta a &= J_2R^2a_0^2p_0^{-3}[\{1+e_0\cos(\omega_0+\Delta\omega)\}^3-\{1+e_0\cos\omega_0\}^3]+\ldots\\
&= -3J_2R^2a^2p^{-3}e\sin\omega(1+e\cos\omega)^2\Delta\omega+O(J_2^3,J_3^2,\ldots).
\end{aligned}$$

Finally, substituting the first-order variation of $\Delta\omega$,

$$\Delta a = -18\pi J_2^2 R^4 a^{-3}(1-e^2)^{-5} e \sin\omega (1+e\cos\omega)^2 \left(1-\frac{5}{4}f\right) + O(J_2^3, J_3^2, \ldots). \quad (54)$$

The following formulae give the secular and long-period perturbations in e, i, Ω, ω, σ^*, ϕ and p, including all terms in J_2, J_3, J_4, J_5, J_6 and J_2^2. To avoid a superfluity of zero suffices, these have been omitted. *All* elements on the right can be assumed to have their ascending node values. (This also applies to equation (54) above.) As before, f stands for $\sin^2 i$.

$$\Delta e = 2\pi[\sum_n J_n (R/p)^n e_n + J_2^2 (R/p)^4 e_{22}] \quad (55)$$

where

$$e_2 = 0$$

$$e_3 = -\frac{3}{2}(1-e^2)\sin i \cos\omega \left(1-\frac{5}{4}f\right)$$

$$e_4 = -\frac{45}{16}(1-e^2)\left(1-\frac{7}{6}f\right) f e \sin 2\omega$$

$$e_5 = \frac{15}{4}(1-e^2)\sin i \left[\left(1-\frac{7}{2}f+\frac{21}{8}f^2\right)\left(1+\frac{3}{4}e^2\right)\cos\omega + \frac{7}{8}\left(1-\frac{9}{8}f\right) f e^2 \cos 3\omega\right]$$

$$e_6 = \frac{525}{32}(1-e^2) f \left[\left(1-3f+\frac{33}{16}f^2\right)\left(1+\frac{1}{2}e^2\right) e \sin 2\omega + \frac{3}{16}\left(1-\frac{11}{10}f\right) f e^3 \sin 4\omega\right]$$

$$e_{22} = -3\sin\omega \left[\frac{3}{2}\left(1-\frac{5}{4}f\right)(1+e\cos\omega)^2 - f(1-e^2)\left(\left(1-\frac{5}{4}f\right) - \left(\frac{7}{8}-\frac{15}{16}f\right)\right) e\cos\omega\right].$$

$$\Delta i = 2\pi[\sum_n J_n (R/p)^n i_n + J_2^2 (R/p)^4 i_{22}] \quad (56)$$

where

$$i_2 = 0$$

$$i_3 = \frac{3}{2}\left(1-\frac{5}{4}f\right) e \cos\omega \cos i$$

$$i_4 = \frac{45}{32}\left(1-\frac{7}{6}f\right) e^2 \sin 2\omega \sin 2i$$

$$i_5 = -\frac{15}{4} e \cos i \left[\left(1-\frac{7}{2}f+\frac{21}{8}f^2\right)\left(1+\frac{3}{4}e^2\right)\cos\omega + \frac{7}{8}\left(1-\frac{9}{8}f\right) f e^2 \cos 3\omega\right]$$

$$i_6 = -\frac{525}{64} e \sin 2i \left[\left(1-3f+\frac{33}{16}f^2\right)\left(1+\frac{1}{2}e^2\right) e \sin 2\omega + \frac{3}{16}\left(1-\frac{11}{10}f\right) f e^3 \sin 4\omega\right]$$

$$i_{22} = -\frac{3}{2}\sin 2i \left[\left(1-\frac{5}{4}f\right) e \sin\omega + \left(-\frac{7}{16}+\frac{15}{32}f\right) e^2 \sin 2\omega\right].$$

$$\Delta\Omega = 2\pi\left[\sum_n J_n(R/p)^n\Omega_n + J_2^2(R/p)^4\Omega_{22}\right] \tag{57}$$

where

$$\Omega_2 = -\frac{3}{2}\cos i$$

$$\Omega_3 = \frac{3}{2}\left(1-\frac{15}{4}f\right)e\sin\omega\cot i$$

$$\Omega_4 = \frac{15}{4}\cos i\left[\left(1-\frac{7}{4}f\right)\left(1+\frac{3}{2}e^2\right)-\frac{3}{4}\left(1-\frac{7}{3}f\right)e^2\cos 2\omega\right]$$

$$\Omega_5 = -\frac{15}{4}\cot i\left[\left(1-\frac{21}{2}f+\frac{105}{8}f^2\right)\left(1+\frac{3}{4}e^2\right)e\sin\omega+\frac{7}{8}\left(1-\frac{15}{8}f\right)e^3\sin 3\omega\right]$$

$$\Omega_6 = -\frac{105}{16}\cos i\left[\left(1-\frac{9}{2}f+\frac{33}{8}f^2\right)\left(1+5e^2+\frac{15}{8}e^4\right)-\right.$$

$$\left.-\frac{5}{2}\left(1-6f+\frac{99}{16}f^2\right)\left(1+\frac{1}{2}e^2\right)e^2\cos 2\omega-\frac{15}{32}\left(1-\frac{33}{20}f\right)fe^4\cos 4\omega\right]$$

$$\Omega_{22} = \frac{3}{2}\cos i\left[\left(\frac{3}{4}-5f\right)+(4-10f)e\cos\omega+\left(-\frac{1}{4}-\frac{5}{16}f\right)e^2+\right.$$

$$\left.+\left(-\frac{7}{8}+\frac{15}{8}f\right)e^2\cos 2\omega\right].$$

$$\Delta\omega = 2\pi\left[\sum_n J_n(R/p)^n\omega_n + J_2^2(R/p)^4\omega_{22}\right] \tag{58}$$

where

$$\omega_2 = 3\left(1-\frac{5}{4}f\right)$$

$$\omega_3 = \frac{3}{2}e^{-1}\sin\omega\sin i\left[\left(1-\frac{5}{4}f\right)+\left(\frac{35}{4}\cos^2 i-\operatorname{cosec}^2 i\right)e^2\right]$$

$$\omega_4 = -\frac{15}{32}\left[(16-62f+49f^2)+(6f-7f^2)\cos 2\omega+\left(18-63f+\frac{189}{4}f^2\right)e^2+\right.$$

$$\left.+\left(-6+35f-\frac{63}{2}f^2\right)e^2\cos 2\omega\right]$$

$$\omega_5 = \frac{105}{16}e^{-1}\sin\omega\operatorname{cosec} i\left[\left(-\frac{4}{7}+2f-\frac{3}{2}f^2\right)f+\left(\frac{4}{7}-\frac{87}{7}f+\frac{67}{2}f^2-\frac{357}{16}f^3\right)e^2+\right.$$

$$+\left(-1+\frac{9}{8}f\right)f^2e^2\cos 2\omega+\left(\frac{3}{7}-7f+\frac{267}{16}f^2-\frac{165}{16}f^3\right)e^4+$$

$$\left.+\left(1-\frac{39}{8}f+\frac{33}{8}f^2\right)fe^4\cos 2\omega\right]$$

$$\omega_6 = \frac{525}{64}\left[\frac{8}{5}\left(1-8f+\frac{129}{8}f^2-\frac{297}{32}f^3\right)+\right.$$
$$+\left(2-6f+\frac{33}{8}f^2\right)f\cos 2\omega+6\left(1-\frac{43}{6}f+\frac{109}{8}f^2-\frac{121}{8}f^3\right)e^2+$$
$$+\left(-2+25f-\frac{459}{8}f^2+\frac{561}{16}f^3\right)e^2\cos 2\omega+\frac{3}{8}\left(1-\frac{11}{10}f\right)f^2e^2\cos 4\omega+$$
$$+\left(2-\frac{27}{2}f+\frac{99}{4}f^2-\frac{429}{32}f^3\right)e^4+\left(-1+\frac{21}{2}f-\frac{363}{16}f^2+\frac{429}{32}f^3\right)e^4\cos 2\omega+$$
$$\left.+\frac{3}{8}\left(-1+\frac{22}{5}f-\frac{143}{40}f^2\right)fe^4\cos 4\omega\right]$$

$$\omega_{22} = \frac{9}{4}\left[\left(-2+\frac{23}{6}f-\frac{5}{3}f^2\right)e^{-1}\cos\omega+\left(\frac{95}{12}f-\frac{445}{48}f^2\right)+\left(-2+\frac{23}{12}f+\frac{5}{8}f^2\right)\cos 2\omega+\right.$$
$$+\left(-\frac{25}{6}+\frac{461}{24}f-\frac{50}{3}f^2\right)e\cos\omega+\left(-\frac{1}{2}+\frac{5}{8}f\right)e\cos 3\omega+$$
$$\left.+\left(\frac{7}{12}-\frac{3}{8}f-\frac{15}{32}f^2\right)e^2+\left(\frac{7}{12}-\frac{79}{24}f+\frac{45}{16}f^2\right)e^2\cos 2\omega\right].$$

$$\Delta\sigma^* = 2\pi\left[\sum_n J_n(R/p)^n\sigma^*_n+J_2^2(R/p)^4\sigma^*_{22}\right] \tag{59}$$

where

$$\sigma^*_2 = \frac{3}{2}(1-e^2)^{\frac{1}{2}}\left(1-\frac{3}{2}f\right)$$

$$\sigma^*_3 = -\frac{3}{2}(1-e^2)^{\frac{1}{2}}(1-4e^2)\left(1-\frac{5}{4}f\right)e^{-1}\sin\omega\sin i$$

$$\sigma^*_4 = \frac{45}{16}(1-e^2)^{\frac{1}{2}}\left[\left(1-\frac{5}{2}e^2\right)\left(1-\frac{7}{6}f\right)f\cos 2\omega+\left(-1+5f-\frac{35}{8}f^2\right)e^2\right]$$

$$\sigma^*_5 = \frac{15}{4}(1-e^2)^{\frac{1}{2}}e^{-1}\sin i\left[\left(1-\frac{7}{2}f+\frac{21}{8}f^2\right)\left(1-\frac{7}{4}e^2-\frac{9}{2}e^4\right)\sin\omega+\right.$$
$$\left.+\frac{7}{8}\left(1-\frac{9}{8}f\right)(1-2e^2)fe^2\sin 3\omega\right]$$

$$\sigma^*_6 = -\frac{35}{16}(1-e^2)^{\frac{1}{2}}\left[\left(1-\frac{21}{2}f+\frac{189}{8}f^2-\frac{231}{16}f^3\right)\left(1-\frac{5}{2}e^2-\frac{15}{8}e^4\right)+\right.$$
$$+\frac{15}{2}\left(1-3f+\frac{33}{16}f^2\right)\left(1-\frac{3}{2}e^2-\frac{7}{4}e^4\right)f\cos 2\omega+$$
$$\left.+\frac{45}{32}\left(1-\frac{11}{10}f\right)\left(1-\frac{7}{4}e^2\right)f^2e^2\cos 4\omega\right]$$

$$\sigma^*_{22} = \frac{9}{2}(1-e^2)^{-\frac{1}{2}}\left[\left(1-\frac{23}{12}f+\frac{5}{6}f^2\right)e^{-1}\cos\omega+\left(\frac{25}{12}f-\frac{131}{48}f^2\right)+\right.$$
$$+\left(1-\frac{23}{24}f-\frac{5}{16}f^2\right)\cos 2\omega+\left(-\frac{7}{4}+\frac{355}{48}f-\frac{20}{3}f^2\right)e\cos\omega+$$
$$+\left(\frac{1}{4}-\frac{5}{16}f\right)e\cos 3\omega+\left(\frac{5}{12}-\frac{49}{24}f+\frac{67}{48}f^2\right)e^2+$$
$$+\left(-\frac{3}{-4}-\frac{23}{24}f+\frac{17}{8}f^2\right)e^2\cos 2\omega+\left(\frac{11}{8}-\frac{77}{12}f+\frac{35}{6}f^2\right)e^3\cos\omega+$$
$$+\left(-\frac{5}{24}+\frac{1}{4}f\right)e^3\cos 3\omega+\left(-\frac{35}{96}+\frac{41}{96}f+\frac{79}{768}f^2\right)e^4+$$
$$\left.+\left(\frac{53}{48}f-\frac{37}{32}f^2\right)e^4\cos 2\omega-\frac{3}{512}f^2e^4\cos 4\omega\right].$$

$$\Delta\phi = 2\pi\left[\sum_n J_n(R/p)^n\phi_n + J_2^2(R/p)^4\phi_{22}\right] \tag{60}$$

where

$$\phi_2 = 3(1-e^2)^{\frac{1}{2}}\left(\frac{3}{2}f-1\right)$$

$$\phi_3 = 12(1-e^2)^{\frac{1}{2}}\left(\frac{5}{4}f-1\right)e\sin\omega\sin i$$

$$\phi_4 = \frac{15}{4}(1-e^2)^{\frac{1}{2}}\left[\left(1-5f+\frac{35}{8}f^2\right)\left(1+\frac{3}{2}e^2\right)+\frac{15}{4}\left(1-\frac{7}{6}f\right)fe^2\cos 2\omega\right]$$

$$\phi_5 = 45(1-e^2)^{\frac{1}{2}}\sin i\left[\left(1-\frac{7}{2}f+\frac{21}{8}f^2\right)\left(1+\frac{3}{4}e^2\right)e\sin\omega+\frac{7}{24}\left(1-\frac{9}{8}f\right)fe^3\sin 3\omega\right]$$

$$\phi_6 = -\frac{35}{8}(1-e^2)^{\frac{1}{2}}\left[\left(1-\frac{21}{2}f+\frac{189}{8}f^2-\frac{231}{16}f^3\right)\left(1+5e^2+\frac{15}{8}e^4\right)+\right.$$
$$\left.+\frac{105}{4}\left(1-3f+\frac{33}{16}f^2\right)\left(1+\frac{1}{2}e^2\right)fe^2\cos 2\omega+\frac{315}{128}\left(1-\frac{11}{10}f\right)f^2e^4\cos 4\omega\right]$$

$$\phi_{22} = \frac{9}{2}(1-e^2)^{\frac{1}{2}}\left[\left(-\frac{9}{2}f+\frac{49}{8}f^2\right)+(1-9f+10f^2)e\cos\omega+\left(\frac{9}{4}f-\frac{21}{16}f^2\right)e^2+\right.$$
$$+\left(-\frac{1}{2}+4f-\frac{33}{8}f^2\right)e^2\cos 2\omega+\left(-\frac{9}{4}+\frac{87}{8}f-10f^2\right)e^3\cos\omega+$$
$$+\left(-\frac{1}{12}+\frac{1}{8}f\right)e^3\cos 3\omega+\left(\frac{5}{8}-\frac{7}{8}f+\frac{7}{64}f^2\right)e^4+$$
$$\left.+\left(-\frac{3}{2}f+\frac{3}{2}f^2\right)e^4\cos 2\omega+\frac{3}{128}f^2e^4\cos 4\omega\right].$$

$$\Delta p = 2\pi\left[\sum_n J_n(R/p)^n p_n + J_2^2(R/p)^4 p_{22}\right] \tag{61}$$

where

$$p_2 = 0$$

$$p_3 = 3p\left(1-\frac{5}{4}f\right)e\cos\omega\sin i$$

$$p_4 = \frac{45}{8}p\left(1-\frac{7}{6}f\right)fe^2\sin 2\omega$$

$$p_5 = -\frac{15}{2}pe\sin i\left[\left(1-\frac{7}{2}f+\frac{21}{8}f^2\right)\left(1+\frac{3}{4}e^2\right)\cos\omega+\frac{7}{8}\left(1-\frac{9}{8}f\right)fe^2\cos 3\omega\right]$$

$$p_6 = -\frac{525}{16}pef\left[\left(1-3f+\frac{33}{16}f^2\right)\left(1+\frac{1}{2}e^2\right)e\sin 2\omega+\frac{3}{16}\left(1-\frac{11}{10}f\right)fe^3\sin 4\omega\right]$$

$$p_{22} = -6pef\left[\left(1-\frac{5}{4}f\right)\sin\omega-\left(\frac{7}{16}-\frac{15}{32}f\right)e\sin 2\omega\right].$$

Checks for the first-order terms, similar to (43) and (44) are:

$$p_n = 2pi_n\tan i = -2pe\,e_n/(1-e^2) \tag{62}$$

and

$$(1-e^2)^{-\frac{1}{2}}(\sigma^*{}_n+\phi_n) = -(\omega_n+\Omega_n\cos i). \tag{63}$$

Some of the above formulae have been obtained by other workers: in particular, Groves (1960) gives explicitly the J_2, J_3, J_4 terms in a, e, i, Ω, ω and σ^* and gives relationships from which the higher first order terms can fairly easily be derived. Zhongolovitch (1960) gives the J_2, J_4 and J_2^2 terms for Ω and ω. In comparing their results with the present ones it should be noted that Groves uses $A_l{}^0$ in place of $-\mu J_l/R$ and Zhongolovitch retains the notation of Jeffreys (1952) with $J = 3J_2/2$ and $D = -35J_4/8$. In all cases the results agree.

10. The nodal period T_Ω

The time t is a more convenient practical variable to use than u, at least when secular changes are being considered. In order to be able to calculate the secular changes as functions of time it is necessary to know T_Ω, i.e. the change in t as u goes from 0 to 2π. The usual procedure for determining T_Ω (e.g. Zhongolovitch 1960) is to expand dt/du as a function of u and integrate term-by-term. Unfortunately, however, there are powers of $(1+e\cos v)$ occurring in the denominators of these terms so that a new series of basic integrals has to be evaluated. It is possible to avoid this by a method which will now be described, in which the essential step is the use of the energy integral (30) to provide a constant differing from the mean motion n by a quantity of order J_2.

Since $n^2a^3 = \mu$ and $(\mu/2a)+F =$ constant, then

$$n\left(1+\frac{2aF}{\mu}\right)^{\frac{1}{2}} = \text{constant} = n' \text{ say.} \tag{64}$$

Now starting from equation (10),

$$\Delta M = \Delta\sigma^* + \int_0^{T_\Omega} n\,dt$$

$$= \Delta\sigma^* + n'T_\Omega - \int_0^{T_\Omega} (n'-n)\,dt$$

$$= \Delta\sigma^* + n'T_\Omega - \int_0^{2\pi} (n'-n)\frac{dt}{du}\,du,$$

so that

$$T_\Omega = \frac{1}{n'}\left[\Delta M - \Delta\sigma^* + \int_0^{2\pi} (n'-n)\frac{dt}{du}\,du\right]. \tag{65}$$

Here, $\Delta\sigma^*$ is known, ΔM can be found by varying the equations $M = E - e\sin E$ and $\cos E = (e+\cos v)/(1+e\cos v)$ and the integral can fairly readily be evaluated.

If the analysis is restricted to first-order perturbations + J_2^2 it is found that

$$\begin{aligned}\Delta M = {} & 2\pi - (1+e_0\cos\omega_0)^{-2}(1-e_0^2)^{\frac{1}{2}}\Delta\omega\{1+(1+e_0\cos\omega_0)^{-1}e_0\sin\omega_0\Delta\omega\} + \\ & + \{(1+e_0\cos\omega_0)^{-1} + (1+e_0\cos\omega_0)^{-2}\}(1-e_0^2)^{\frac{1}{2}}\Delta e\sin\omega_0,\end{aligned} \tag{66}$$

which can be calculated using the expressions (58), (55) for $\Delta\omega$, Δe.

To determine the integral in (65) it is seen first from (64) that

$$n'-n = 3na\frac{F}{\mu} + \frac{3}{2}na^2\frac{F^2}{\mu^2} + O(F^3), \tag{67}$$

so that

$$\int_0^{2\pi} (n'-n)\frac{dt}{du}\,du = \int_0^{2\pi} \left\{3na\frac{F}{\mu} + \frac{3}{2}na^2\frac{F^2}{\mu^2}\right\}\frac{dt}{du}\,du + O(F^3). \tag{68}$$

The first-order part of this integral can be evaluated by using the auxiliary element ϕ.

If the expressions for f_1 (25) and F (28) are compared, it will be seen that the u-dependent parts in each term are similar: and if we let

$$f_1 = \sum_l f_{1l}J_l,$$

then

$$F = -\sum r f_{1l}J_l/(l+1). \tag{69}$$

Now, from (21),

$$\begin{aligned}\dot{\phi} &= 2nar\mu^{-1}f_1 \\ &= 2nar\mu^{-1}\sum_l f_{1l}J_l,\end{aligned}$$

and so

$$\Delta\phi = \int_0^{2\pi} 2nar\mu^{-1}\sum_l f_{1l}J_l \cdot \frac{dt}{du}\,du.$$

Comparing this with (60) gives

$$\left.\begin{aligned}&\int_0^{2\pi} 2nar\mu^{-1}f_{12}J_2\frac{dt}{du}\,du = 2\pi[J_2(R/p_0)^2\phi_2 + J_2^2(R/p_0)^4\phi_{22}] \\ \text{and}\quad &\int_0^{2\pi} 2nar\mu^{-1}f_{1l}J_l\frac{dt}{du}\,du = 2\pi J_l(R/p_0)^2\phi_l, \qquad l > 2.\end{aligned}\right\} \tag{70}$$

Thus, using (69) and (70),

$$\begin{aligned}\int_0^{2\pi} 3na\frac{F}{\mu}\cdot\frac{dt}{du}\,du &= -\frac{3}{2}\sum_l \frac{1}{l+1}\int_0^{2\pi} 2nar\mu^{-1}J_lf_{1l}\cdot\frac{dt}{du}\,du \\ &= -2\pi\left[\sum_l J_l(R/p_0)^l\cdot\frac{3\phi_l}{2(l+1)} + \frac{1}{2}J_2^2(R/p_0)^4\phi_{22}\right]. \end{aligned} \tag{71}$$

Finally, the remaining integral in (68) is

$$\begin{aligned}\int_0^{2\pi}\frac{3}{2}na^2\frac{F^2}{\mu^2}\cdot\frac{dt}{du}\,du &= \frac{3}{8}J_2^2(R/p)^4(1-e^2)^{-\frac{1}{2}}\int_0^{2\pi}(3fS^2-1)^2(1+e\cos v)^4\,du + \\ &\quad + O(J_2^3, J_3^2, \ldots) \\ &= \frac{3}{4}\pi J_2^2(R/p_0)^4(1-e_0^2)^{-\frac{1}{2}}\left[\left(1-3f_0+\frac{27}{8}f_0^2\right)\left(1+3e_0^2+\frac{3}{8}e_0^4\right)+\right. \\ &\quad + \frac{9}{2}\left(1-\frac{3}{2}f_0\right)\left(1+\frac{1}{6}e_0^2\right)f_0e_0^2\cos 2\omega_0 + \\ &\quad \left. + \frac{9}{128}f_0^2e_0^4\cos 4\omega_0\right] + O(J_2^3, J_3^2, \ldots).\end{aligned} \tag{72}$$

This completes the evaluation of the components of T_Ω and it remains to express n' in terms of osculating elements.

From (67), setting $u = 0$,

$$n' = n_0\left(1 + 3a_0\mu^{-1}F_0 + \frac{3}{2}a_0^2\mu^{-2}F_0^2\right) + O(F^3),$$

and from (28),

$$a_0\mu^{-1}F_0 = \frac{1}{2}J_2R^2a_0r_0^{-3} - \frac{3}{8}J_4R^4a_0r_0^{-5} + \frac{5}{16}J_6R^6a_0r_0^{-7} + O(J_8).$$

Thus, setting

$$p_0/r_0 = 1 + e_0\cos\omega_0, \quad \text{and} \quad a_0 = p_0(1-e_0^2)^{-1},$$

$$\begin{aligned} n' = n_0\Big[1 &+ \frac{3}{2}J_2(R/p_0)^2(1-e_0^2)^{-1}(1+e_0\cos\omega_0)^3 - \frac{9}{8}J_4(R/p_0)^4(1-e_0^2)^{-1}\times \\ &\times(1+e_0\cos\omega_0)^5 + \frac{15}{16}J_6(R/p_0)^6(1-e_0^2)^{-1}(1+e_0\cos\omega_0)^7 + \\ &+ \frac{3}{8}J_2^2(R/p_0)^4(1-e_0^2)^{-2}(1+e_0\cos\omega_0)^6 + \ldots\Big]. \end{aligned} \tag{73}$$

Adding together the various components of T_Ω gives finally

$$\begin{aligned} T_\Omega = \frac{2\pi}{n'}\Big[1 &- (r_0/p_0)^2(1-e_0^2)^{\frac{1}{2}}\left(\frac{\Delta\omega}{2\pi}\right)\{1 + (r_0/p_0)e_0\sin\omega_0\Delta\omega\} + \\ &+ \{(r_0/p_0) + (r_0/p_0)^2\}(1-e_0^2)^{\frac{1}{2}}\sin\omega_0 \cdot \frac{\Delta e}{2\pi} - \frac{\Delta\sigma^*}{2\pi} - \\ &- \sum_{l=2}^{\infty} J_l(R/p_0)^l \cdot 3\phi_l/2(l+1) - \frac{1}{2}J_2^2(R/p_0)^4\phi_{22} + \\ &+ \frac{3}{8}J_2^2(R/p_0)^4(1-e_0^2)^{-\frac{1}{2}}\Big\{\left(1 - 3f_0 + \frac{27}{8}f_0^2\right)\left(1 + 3e_0^2 + \frac{3}{8}e_0^4\right) + \\ &+ \frac{9}{2}\left(1 - \frac{3}{2}f_0\right)\left(1 + \frac{1}{6}e_0^2\right)f_0e_0^2\cos 2\omega_0 + \frac{9}{128}f_0^2e_0^4\cos 4\omega_0\Big\} + \\ &+ O(J_2^3, J_3^2, \ldots)\Big], \end{aligned} \tag{74}$$

where n' is given by (73), and $\Delta\omega$, Δe, $\Delta\sigma^*$, ϕ_l by (58), (55), (59), (60).

For many purposes the first-order J_2 term in T_Ω is adequate, and to this order

$$\begin{aligned} T_\Omega = \frac{2\pi}{n_0}\Big[1 - 3J_2(R/p)^2(1+e\cos\omega)^{-2}(1-e^2)^{\frac{1}{2}}\left(1 - \frac{5}{4}\sin^2 i\right)\Big]\times \\ \times\left[1 + \frac{3}{2}J_2(R/p)^2(1-e^2)^{-1}(1+e\cos\omega)^3\right]^{-1}. \end{aligned} \tag{75}$$

11. Transformation of the elements

It has already been explained that the use of osculating elements in the reduction and subsequent analysis of Earth-satellite observations is not very satisfactory if great accuracy is required because the amplitudes of the periodic oscillations in the elements are unduly high. This drawback is a direct result of forcing the orbital plane to follow the oscillations of the satellite's velocity vector. As will be shown in this and the following sections, it is possible to define a set of elements which are much smoother in their behaviour than the osculating elements.

Since the short-term periodic oscillations have only been determined for the J_2 term in the potential, it will be assumed for the present purposes that J_3 and higher coefficients are all zero. The effects of the higher coefficients can be added later.

In this treatment the new set of elements is assumed to differ from osculating elements by quantities of order J_2. An orbital plane, in which the satellite lies, is effectively defined by a new inclination i' and right ascension Ω', and a reference ellipse, with latus rectum p', eccentricity e', is introduced. The satellite departs from this ellipse by quantities of order J_2.

Suppose then that

$$\left.\begin{aligned} u' &= u + J_2\gamma_u \\ v' &= v + J_2\gamma_v \end{aligned}\right\} \tag{76}$$

and for each element ζ,

$$\zeta' = \zeta + J_2\gamma_\zeta\,, \tag{77}$$

except that, for convenience in writing,

$$\sigma' = \sigma^* + J_2\gamma_\sigma\,. \tag{78}$$

Then the geometrical and other assumptions made are contained in the following formulae:

Osculating System	*Smoothed System*	
$\sin\theta\cos(\lambda-\Omega) = \cos u$	$\sin\theta\cos(\lambda-\Omega') = \cos u'$	(79)
$\sin\theta\sin(\lambda-\Omega) = \cos i\sin u$	$\sin\theta\sin(\lambda-\Omega') = \cos i'\sin u'$	(80)
$\cos\theta = \sin i\sin u$	$\cos\theta = \sin i'\sin u'$	(81)
$\cos E = (e+\cos v)/(1+e\cos v)$	$\cos E' = (e'+\cos v')/(1+e'\cos v')$	(82)
$\sin E = (1-e^2)^{\frac{1}{2}}\sin v/(1+e\cos v)$	$\sin E' = (1-e'^2)^{\frac{1}{2}}\sin v'/(1+e'\cos v')$	(83)
$p/r = 1+e\cos v$	$p'/r = 1+e'\cos v'+J_2 z$	(84)
$E-e\sin E = \int_0^t n\,dt+\sigma^*$	$E'-e'\sin E' = n't+\sigma'$	(85)
$u = v+\omega$	$u' = v'+\omega'$	(86)

In addition, n' is assumed to be the fundamental constant already defined (equation (64)) in Section 10.

In Section 12 attention is restricted to first-order solutions and the first step is to find the relationships between the γ-functions. These relationships reduce

to two basic equations which are then separated into further equations by equating coefficients of e. At this stage it is found that, apart from minor details, there is an obvious solution of great simplicity which is then studied further. Taking this simple first-order solution as a basis, a similar procedure is adopted in Section 13 to determine a second-order solution.

12. Relationships between the γ-functions

From equations (79), by subtraction, substitution for Ω', u' from (77), (76) and expansion in powers of J_2,

$$\gamma_u = -\cos i\,\gamma_\Omega + \frac{1}{2}J_2 \cot u(\gamma_\Omega{}^2 - \gamma_u{}^2) + O(J_2{}^2). \tag{87}$$

Similarly, from equations (81),

$$\gamma_i = -\tan i \cot u\,\gamma_u - J_2 \cot u\,\gamma_i\gamma_u + \frac{1}{2}J_2 \tan i(\gamma_i{}^2 + \gamma_u{}^2) + O(J_2{}^2). \tag{88}$$

From (86) and (87),

$$\begin{aligned}\gamma_v &= \gamma_u - \gamma_\omega \\ &= -(\gamma_\Omega \cos i + \gamma_\omega) + O(J_2).\end{aligned} \tag{89}$$

From (82),

$$\begin{aligned}&-\gamma_E \sin E(1+e\cos v) - \gamma_e(1-\cos E\cos v) + \gamma_v \sin v(1-e\cos E) \\ &\quad = \frac{1}{2}J_2\gamma_E{}^2 \cos E(1+e\cos v) - \frac{1}{2}J_2\gamma_e{}^2 \cos v(1-e\cos E) + J_2\gamma_E\gamma_e \sin E\cos v + \\ &\qquad + J_2\gamma_v\gamma_e \sin v\cos E - J_2\gamma_E\gamma_v e\sin v\sin E + O(J_2{}^2),\end{aligned} \tag{90}$$

or, eliminating E, with the aid of (82), (83) and (84),

$$\gamma_E = \frac{r}{p}(1-e^2)^{\frac{1}{2}}\gamma_v - \frac{r}{p}(1-e^2)^{-\frac{1}{2}} \sin v\,\gamma_e + O(J_2). \tag{91}$$

From equations (85),

$$\begin{aligned}&J_2\gamma_E(1-e\cos E) - J_2\gamma_e \sin E + \frac{1}{2}J_2{}^2\gamma_E{}^2 e\sin E - J_2{}^2\gamma_E\gamma_e \cos E \\ &\qquad = \int_0^t (n'-n)\,dt + J_2\gamma_\sigma + O(J_2{}^3).\end{aligned} \tag{92}$$

Now, by (67),

$$n'-n = 3na\frac{F}{\mu} + \frac{3}{2}na^2\frac{F^2}{\mu^2},$$

and since the J_2 terms alone are being considered,

$$F = -\frac{1}{3}rf_1,$$

so that

$$n'-n = -nar\frac{f_1}{\mu}+\frac{na^2f_1^2r^2}{\mu^2}$$

$$= -\tfrac{1}{2}\phi+na^2f_1^2r^2/\mu^2. \tag{93}$$

Substituting in (92) and eliminating E, with the aid of (82), (83), (84),

$$\gamma_E(1-e^2)(r/p)-\gamma_e(1-e^2)^{\frac{1}{2}}(r/p)\sin v = -\tfrac{1}{2}J_2^{-1}\delta\phi+\gamma_\sigma+O(J_2). \tag{94}$$

Finally, eliminating γ_E between (91) and (94),

$$(1-e^2)\gamma_v-\gamma_e\sin v\{1+(p/r)\} = (1-e^2)^{-1}(p/r)^2(\gamma_\sigma-\tfrac{1}{2}J_2^{-1}\delta\phi+)O(J_2). \tag{95}$$

The remaining relationship between the γ's comes from equations (84), which give

$$z = (\gamma_p/r)+\gamma_e e\sin v-\gamma_e\cos v+J_2\gamma_e\gamma_v\sin v+\tfrac{1}{2}J_2\gamma_v^2e\cos v+O(J_2^2). \tag{96}$$

We have now to consider possible solutions of the equations (87), (88), (89), (95) and (96). Bearing in mind that

$$\zeta' = \zeta+J_2\gamma_\zeta$$

and

$$\zeta = \zeta_0+\delta\zeta,$$

what we would like is for

$$J_2\gamma_\zeta = J_2c_\zeta-\delta\zeta_{\text{per}},$$

where $\delta\zeta_{\text{per}}$ is the periodic component of $\delta\zeta$ and c_ζ = constant $+O(J_2)$, for then the periodic component would disappear from the dashed elements, leaving only secular components.

Of course this cannot be done completely, but the idea suggests the method of procedure to be adopted.

Let

$$J_2\gamma_\zeta = J_2c_\zeta-\delta\zeta_{\text{per}} \tag{97}$$

and

$$\delta\zeta = \delta\zeta_{\text{sec}}+\delta\zeta_{\text{per}} \tag{98}$$

so that

$$\zeta' = \zeta_0+\delta\zeta_{\text{sec}}+J_2c_\zeta. \tag{99}$$

Thus, after substituting for the γ's from (97) in the equations relating the γ's, equations for the c's are obtained which are solved in such a way as to keep the periodic content as small as possible.

The two basic equations are (95) and (96) which yield c_v, c_e, c_σ, c_p and z.

Having solved these, equations (87), (88) and (89) give c_Ω, c_ω and c_i.

Firstly, from (95), using (97) and (99),

$$(1-e^2)c_v-\sin v(1+p/r)c_e-(1-e^2)^{-\frac{1}{2}}(p/r)^2c_\sigma$$

$$= -J_2^{-1}(1-e^2)^{-\frac{1}{2}}(p/r)^2(\delta\sigma^*_{\rm per}+\tfrac{1}{2}\delta\phi)-$$

$$-J_2^{-1}\sin v(1+p/r)\delta e_{\rm per}-J_2^{-1}(1-e^2)(\delta\Omega_{\rm per}\cos i+\delta\omega_{\rm per})+O(J_2). \quad (100)$$

Secondly, from (96), using (97) and (89),

$$z-c_p/r-c_v e\sin v+c_e\cos v = -J_2^{-1}\delta p_{\rm per}+J_2^{-1}\delta e_{\rm per}\cos v+J_2^{-1}e\sin v\times$$
$$\times(\delta\Omega_{\rm per}\cos i+\delta\omega_{\rm per})+O(J_2). \quad (101)$$

The right-hand sides of these equations are known functions which can be calculated using equations (36) to (42). As these functions are too complicated to be digested *en masse*, it is convenient to write the c's as polynomials in e and to equate coefficients of equal powers of e.

There is a secular term on the right-hand side of (100), from $\delta\phi$, which it is convenient to associate with c_σ.

Let

$$\left.\begin{aligned} c_v &= -3(R/p^2)(c_{v,-1}e^{-1}+c_{v0}+c_{v1}e+c_{v2}e^2+\ldots)\\ c_e &= 3(R/p)^2(c_{e,-1}+c_{e0}+c_{e1}e+c_{e2}e^2+\ldots)\\ c_p &= 3p(R/p)^2(c_{p,-1}e^{-1}+c_{p0}+c_{p1}e+c_{p2}e^2+\ldots)\\ (1-e^2)^{-\frac{1}{2}}(c_\sigma-\tfrac{1}{2}J_2^{-1}\delta\phi_{\rm sec}) &= 3(R/p)^2(c_{\sigma,-1}e^{-1}+c_{\sigma0}+c_{\sigma1}e+c_{\sigma2}e^2+\ldots)\end{aligned}\right\} \quad (102)$$

and

$$z = 3(R/p)^2(z_{-1}e^{-1}+z_0+z_1e+z_2e^2+\ldots),$$

the factor $3(R/p)^2$ being introduced to simplify the resulting equations.

Substituting these expressions in (100) and (101) and substituting for the δ's from (36) to (42) it is found, after some considerable labour, that

$$\left.\begin{aligned} c_{\sigma,-1}+2c_{e,-1}\sin v+c_{v,-1} &= 0,\\ 2(c_{\sigma,-1}\cos v+c_{e0}\sin v)+c_{e,-1}\sin v\cos v+c_{\sigma0}+c_{v0} &= Q_0,\\ 2(c_{\sigma0}\cos v+c_{e1}\sin v)+\cos v(c_{\sigma,-1}\cos v+c_{e0}\sin v)+c_{\sigma1}+c_{v1}-c_{v,-1} &= Q_1,\\ 2(c_{\sigma1}\cos v+c_{e2}\sin v)+\cos v(c_{\sigma0}\cos v+c_{e1}\sin v)+c_{\sigma2}+c_{v2}-c_{v0} &= Q_2,\\ \ldots\ldots\ldots\

$$Q_1 = -\frac{2}{3}f\sin\omega+\left(-\frac{1}{2}-\frac{1}{2}f\right)\sin v+\left(\frac{1}{6}f-\frac{1}{2}\right)\sin(2u-v)+$$
$$+\left(\frac{1}{3}f-\frac{1}{2}\right)\cos v(\sin v\cos\omega+\cos v\sin\omega),$$

$$Q_2 = \left(\frac{1}{4}-\frac{2}{3}f\right)\sin\omega\cos v+\left(-\frac{1}{4}-\frac{2}{3}f\right)\cos\omega\sin v-\frac{1}{12}\sin(3\omega+v)-$$
$$-\frac{5}{12}\sin v\cos v-\frac{1}{4}\cos v\sin(2u-v)-\frac{1}{4}\sin 2\omega-\frac{1}{24}f\sin 2u,$$

$$Q_3 = \left(\frac{1}{2}f-\frac{1}{3}\right)\sin v+\left(\frac{1}{8}-\frac{1}{3}f\right)\sin\omega\cos^2 v+\left(-\frac{1}{8}-\frac{1}{3}f\right)\cos\omega\sin v\cos v-$$
$$-\frac{1}{24}\cos v\sin(3\omega+v)+\left(\frac{2}{3}f-\frac{3}{8}\right)\sin\omega-\frac{1}{24}\sin 3\omega-\frac{1}{6}f\sin(2u-v),$$

$$Q_4 = Q_5 = \ldots = 0,$$

$$P_0 = \left(\frac{1}{3}f-\frac{1}{2}\right)(\cos v\cos\omega-\sin v\sin\omega)+\left(\frac{1}{2}-\frac{1}{4}f\right)-\frac{1}{12}f\cos 2u,$$

$$P_1 = \frac{2}{3}f\cos\omega-\frac{1}{4}\cos(2u-v)-\frac{1}{6}f\cos 2u\cos v+\left(\frac{1}{4}-\frac{1}{2}f\right)\cos v+$$
$$+\frac{1}{24}f\sin 2u\sin v,$$

$$P_2 = \left(\frac{1}{3}f-\frac{1}{8}\right)\cos\omega\cos v+\left(\frac{3}{8}-\frac{2}{3}f\right)\sin\omega\sin v+\left(\frac{1}{4}-\frac{3}{8}f\right)-\frac{1}{24}\cos(3\omega+v)+$$
$$+\left(\frac{1}{8}f-\frac{1}{12}\right)\cos 2v-\frac{1}{12}f\cos 2u\cos^2 v+\frac{1}{6}f\sin v\sin(2u-v),$$

$$P_3 = P_4 = \ldots = 0.$$

In studying equations (103) and (104) it is noted first that, whatever values are given to the c's, equations (104) can be solved for the x's. Attention will therefore be concentrated on equations (103) to determine the restrictions to be placed on the c's.

It is seen that the c_v's can be eliminated by adding first the odd equations in (103) and then the even equations, so let

$$\left.\begin{aligned} c_{\sigma,-1}+c_{\sigma 1}+c_{\sigma 3}+\ldots &= c_\sigma' \\ c_{\sigma 0}+c_{\sigma 2}+c_{\sigma 4}+\ldots &= c_\sigma'' \\ c_{\varepsilon,-1}+c_{\varepsilon 1}+c_{\varepsilon 3}+\ldots &= c_\varepsilon' \\ c_{\varepsilon 0}+c_{\varepsilon 2}+c_{\varepsilon 4}+\ldots &= c_\varepsilon''. \end{aligned}\right\} \tag{105}$$

Then addition of the odd equations gives

$$(1+\cos^2 v)c_\sigma'+2\cos v\,c_\sigma''+\sin v\cos v\,c_\varepsilon'+2\sin v\,c_\varepsilon'' = Q_1+Q_3, \tag{106}$$

and addition of the even equations gives

$$2\cos v\,c_\sigma'+(1+\cos^2 v)c_\sigma''+2\sin v\,c_\varepsilon'+\sin v\cos v\,c_\varepsilon'' = Q_0+Q_2. \tag{107}$$

In solving these it must be remembered that we require the c-functions to be bounded, i.e. they must not have $\sin v$ or $\cos v$ in their denominators. Eliminating c_σ'' between (106) and (107) gives

$$c_\varepsilon'' \sin v = [(Q_1+Q_3)-(Q_0+Q_2)\cos v]-\frac{1}{2}(Q_1+Q_3)\sin^2 v-\frac{1}{2}c_\sigma' \sin^4 v+$$
$$+\left(\frac{3}{2}-\frac{1}{2}\cos^2 v\right)\sin v\cos v\, c_\varepsilon', \tag{108}$$

so that, for c_ε'' to be bounded it is necessary that $(Q_1+Q_3)-(Q_0+Q_2)\cos v$ be explicitly divisible by $\sin v$.

Now

$$Q_1+Q_3 = -\frac{3}{8}\sin\omega-\frac{3}{8}\sin\omega\cos^2 v-\frac{5}{8}\cos\omega\sin v\cos v-\frac{1}{24}\sin 3\omega-$$
$$-\frac{1}{24}\cos v\sin(3\omega+v)-\frac{1}{2}\sin(2u-v)-\frac{5}{6}\sin v, \tag{109}$$

and

$$Q_0+Q_2 = -\frac{3}{4}\sin\omega\cos v-\frac{5}{4}\cos\omega\sin v-\frac{1}{12}\sin(3\omega+v)-\frac{1}{4}\sin 2u\cos^2 v+$$
$$+\frac{1}{4}\cos 2u\sin v\cos v-\frac{5}{12}\sin v\cos v-\frac{1}{4}\sin 2\omega, \tag{110}$$

whence

$$(Q_1+Q_3)-\cos v(Q_0+Q_2) = \sin v\left[-\frac{3}{8}\sin\omega\sin v+\frac{5}{8}\cos\omega\cos v+\right.$$
$$\left.+\frac{1}{24}\cos(3\omega+v)-\frac{1}{4}\cos 2\omega-\frac{5}{6}+\frac{5}{12}\cos^2 v-\frac{1}{4}\sin v\sin(2u-v)\right]. \tag{111}$$

Then, from (109) and (111),

$$\operatorname{cosec} v\,[(Q_1+Q_3)-\cos v(Q_0+Q_2)]-\frac{1}{2}(Q_1+Q_3)\sin v$$
$$=\left(-\frac{5}{12}-\frac{1}{4}\cos 2\omega\right)+\left(-\frac{3}{16}\sin\omega-\frac{1}{48}\sin 3\omega\right)\sin^2 v+$$
$$+\left(\frac{3}{2}\cos v-\frac{1}{2}\cos^3 v\right)\left(\frac{5}{8}\cos\omega+\frac{1}{24}\sin 3\omega\right),$$

which substituted in (108) gives

$$c_\varepsilon'' = \left(-\frac{5}{12}-\frac{1}{4}\cos 2\omega\right)-\left(\frac{3}{16}\sin\omega+\frac{1}{48}\sin 3\omega+\frac{1}{2}c_\sigma'\right)\sin^2 v+$$
$$+\left(\frac{5}{8}\cos\omega+\frac{1}{24}\sin 3\omega+c_\varepsilon'\right)\left(\frac{3}{2}\cos v-\frac{1}{2}\cos^3 v\right). \tag{112}$$

This equation has the obvious and compelling solution

$$c_e'' = -\frac{5}{12} - \frac{1}{4}\cos 2\omega$$

$$c_\sigma' = -\frac{3}{8}\sin\omega - \frac{1}{24}\sin 3\omega$$

$$c_e' = -\frac{5}{8}\cos\omega - \frac{1}{24}\sin 3\omega$$

where short-term periodic components are absent.

Substituting these values in (106) and solving for c_σ'' gives

$$c_\sigma'' = -\frac{1}{4}\sin 2\omega$$

so that it appears possible to keep e' and σ' constant to the first order in J_2.

Let us therefore make the assumption that the c_σ's and c_e's do not contain u or v explicitly and see where this leads.

Now it is advantageous to have the sequences of c's of finite length, and inspection of (103), (104) shows that this can be done by setting

$$c_{ej} = 0, \quad j \geqslant 4; \qquad c_{\sigma j} = 0, \quad j \geqslant 3; \qquad c_{vj} = 0, \quad j \geqslant 3; \qquad c_{pj} = 0, \quad j \geqslant 3;$$

and

$$z_j = 0, \quad j \geqslant 4.$$

Then it is found from equations (103) that

$$c_{e,-1} = 0,$$

$$c_{e,0} = \left(\frac{1}{3}f - \frac{1}{2}\right)\cos\omega,$$

$$c_{e1} + c_{e3} = -\frac{5}{12} - \frac{1}{4}\cos 2\omega$$

$$c_{e2} = -\frac{1}{6}\cos^3\omega - \frac{1}{3}f\cos\omega,$$

$$c_{\sigma,-1} = \left(\frac{1}{3}f - \frac{1}{2}\right)\sin\omega,$$

$$c_{\sigma 0} + c_{\sigma 2} = -\frac{1}{4}\sin 2\omega,$$

$$c_{\sigma 1} = \frac{1}{6}\sin^3\omega - \frac{1}{3}f\sin\omega,$$

$$c_{v,-1} = \left(\frac{1}{2} - \frac{1}{3}f\right)\sin\omega,$$

$$c_{v0} + c_{\sigma 0} = \frac{1}{24}f\sin 2u,$$

together with the equations

$$2(c_{\sigma 0}\cos v + c_{e1}\sin v) + c_{v1} = \left(\frac{3}{8} - \frac{2}{3}f\right)\sin\omega + \frac{1}{24}\sin 3\omega - \left(\frac{1}{2} + \frac{1}{2}f\right)\sin v + \\ + \left(\frac{1}{6}f - \frac{1}{2}\right)\sin(v+2\omega),$$

$$\cos v(c_{\sigma 0}\cos v + c_{e1}\sin v) + c_{v2} = -\frac{5}{12}\sin v\cos v - \frac{1}{4}\cos v\sin(v+2\omega).$$

The simplest solution of these is

$$c_{e1} = \left(-\frac{1}{4} - \frac{1}{4}f\right) - \frac{1}{4}\cos 2\omega,$$

$$c_{\sigma 0} = -\frac{1}{4}\sin 2\omega,$$

$$c_{v1} = \left(\frac{3}{8} - \frac{2}{3}f\right)\sin\omega + \frac{1}{24}\sin 3\omega + \frac{1}{6}f\sin(v+2\omega),$$

$$c_{v2} = \left(\frac{1}{4}f - \frac{1}{6}\right)\sin v\cos v.$$

Equations (102) then give

$$c_v = -3(R/p)^2\left[\left(\frac{1}{2} - \frac{1}{3}f\right)e^{-1}\sin\omega + \frac{1}{24}f\sin 2u + \frac{1}{4}\sin 2\omega + \left(\frac{3}{8} - \frac{2}{3}f\right)e\sin\omega + \right. \\ \left. + \frac{1}{24}e\sin 3\omega + \frac{1}{6}fe\sin(v+2\omega) + \left(\frac{1}{4}f - \frac{1}{6}\right)e^2\sin v\cos v\right] + O(J_2), \qquad (113)$$

$$c_e = 3(R/p)^2\left[\left(\frac{1}{3}f - \frac{1}{2}\right)\cos\omega - \left(\frac{1}{4} + \frac{1}{4}f\right)e - \frac{1}{4}e\cos 2\omega - \frac{1}{3}fe^2\cos\omega - \right. \\ \left. - \frac{1}{6}e^2\cos^3\omega + \left(\frac{1}{4}f - \frac{1}{6}\right)e^3\right] + O(J_2), \qquad (114)$$

and

$$(1-e^2)^{-\frac{1}{2}}\left(c_\sigma - \frac{1}{2}J_2^{-1}\delta\phi_{\text{sec}}\right) = 3(R/p)^2\left[\left(\frac{1}{3}f - \frac{1}{2}\right)e^{-1}\sin\omega - \frac{1}{4}\sin 2\omega + \right. \\ \left. + \left(\frac{1}{8} - \frac{1}{3}f\right)e\sin\omega - \frac{1}{24}e\sin 3\omega\right] + O(J_2). \qquad (115)$$

Turning now to equations (104) it is seen that, since $c_{e,-1} = 0$, we can take

$$z_{-1} = c_{p,-1} = 0. \qquad (116)$$

The remaining equations become

$$\left.\begin{aligned} z_0 - c_{p0} &= \left(\frac{1}{2} - \frac{1}{4}f\right) - \frac{1}{12}f\cos 2u. \\ z_1 - c_{p1} - c_{p0}\cos v &= \frac{2}{3}f\cos\omega + \left(\frac{1}{2} - \frac{1}{4}f\right)\cos v - \frac{1}{6}f\cos 2u\cos v, \\ z_2 - c_{p2} - c_{p1}\cos v &= \frac{2}{3}f\cos\omega\cos v + \left(\frac{1}{8}f - \frac{1}{12}\right)\cos 2v - \\ &\qquad - \frac{1}{12}f\cos 2u\cos^2 v + \left(\frac{1}{4} - \frac{3}{8}f\right), \\ z_3 - c_{p2}\cos v &= \left(\frac{1}{6} - \frac{1}{4}f\right)\cos v(1+\sin^2 v). \end{aligned}\right\} \tag{117}$$

As there is no obvious solution of these equations it is convenient to impose further restrictions. Firstly, all periodic components in (117) can be ascribed to z, so that c_p (and hence p') will be constant $+ O(J_2^2)$. Secondly, since p' and e' are constant to the first order in J_2, we can define

$$p' = a'(1-e'^2) \tag{118}$$

where a' is given by

$$n'^2 a'^3 = \mu. \tag{119}$$

These conditions are sufficient to define p' and hence c_p and z, for from (118), (119) and setting $e' = e_0 + J_2 c_e$, with c_e given by (114),

$$p' = p_0 - J_2 p(R/p)^2\left[1 + 2fe\cos\omega + \left(1 - \frac{3}{2}f\right)e^2\right] + O(J_2^2). \tag{120}$$

Also, from (99), since $\delta p_{\text{sec}} = O(J_2^2)$,

$$p' = p_0 + J_2 c_p + O(J_2^2),$$

and hence,

$$c_p = -3p(R/p)^2\left[\frac{1}{3} + \frac{2}{3}fe\cos\omega + \left(\frac{1}{3} - \frac{1}{2}f\right)e^2\right] + O(J_2^2), \tag{121}$$

$$z = (p/r)^2\left[\left(\frac{1}{6} - \frac{1}{4}f\right)(1 - e\cos v) - \frac{1}{12}f\cos 2u\right] + O(J_2). \tag{122}$$

This completes the solutions of equations (100) and (101) and it now remains to consider the perturbations in i, Ω and ω.

First, from (87) and (88),

$$\gamma_i = \sin i\cot u\gamma_\Omega + O(J_2), \tag{123}$$

so that, for a non-singular solution, γ_Ω has to be an explicit multiple of $\sin u$.

Now from (97) and (38),

$$\begin{aligned} \gamma_\Omega - c_\Omega &= -J_2^{-1}\delta\Omega_{\text{per}} \\ &= 3(R/p)^2\cos i\left[-\frac{1}{2}SC + \frac{1}{3}S^2 e\sin v + \frac{2}{3}e\sin\omega(1-C)\right] + O(J_2), \end{aligned}$$

or since

$$\sin\omega = S\cos v - C\sin v,$$

$$\gamma_\Omega - c_\Omega = -J_2^{-1}\delta\Omega_{\text{per}} = 3(R/p)^2\cos i\left[-\frac{1}{2}SC-\frac{1}{3}S^2e\sin v-\frac{2}{3}SCe\cos v+\right.$$

$$\left.+\frac{2}{3}e\sin v+\frac{2}{3}e\sin\omega\right]+O(J_2).$$

We therefore take

$$\gamma_\Omega = -3(R/p)^2\cos i\left[\frac{1}{2}SC+\frac{1}{3}S^2e\sin v+\frac{2}{3}SCe\cos v\right]+O(J_2), \tag{124}$$

$$c_\Omega = -3(R/p)^2\cos i\left[\frac{2}{3}e\sin v+\frac{2}{3}e\sin\omega\right]+O(J_2). \tag{125}$$

Then, from (123),

$$\gamma_i = -3(R/p)^2\sin i\cos i\left[\frac{1}{2}C^2+\frac{1}{3}SCe\sin v+\frac{2}{3}C^2e\cos v\right]+O(J_2), \tag{126}$$

so that

$$c_i = -3(R/p)^2\sin i\cos i\left[\frac{1}{2}+\frac{1}{3}e\cos v+\frac{1}{3}e\cos\omega\right]+O(J_2), \tag{127}$$

and from (89), (97), (113) and (125),

$$\begin{aligned} c_\omega &= -c_v - c_\Omega\cos i \\ &= 3(R/p)^2\left[\left(\frac{1}{2}-\frac{1}{3}f\right)e^{-1}\sin\omega+\frac{1}{24}f\sin 2u+\frac{1}{4}\sin 2\omega+\left(\frac{25}{24}-\frac{4}{3}f\right)e\sin\omega+\right. \\ &\left.+\left(\frac{2}{3}-\frac{2}{3}f\right)e\sin v+\frac{1}{24}e\sin 3\omega+\frac{1}{6}fe\sin(2u-v)+\left(\frac{1}{8}f-\frac{1}{12}\right)e^2\sin 2v\right]+O(J_2). \end{aligned} \tag{128}$$

By equation (99), and using the c's, the smoothed elements are expressed in terms of initial osculating elements.

Also setting $u = 0$ in (99), for any element ζ,

$$\zeta_0' = \zeta_0 + J_2c_{\zeta 0}, \tag{129}$$

or subtracting from (99),

$$\zeta' = \zeta_0' + \delta\zeta_{\text{sec}} + J_2(c_\zeta - c_{\zeta 0}). \tag{130}$$

Using this equation, together with the equation for the c's it is found that

$$i' = i_0' + \frac{1}{2}J_2(R/p)^2\sin 2i(e\cos\omega - e\cos v)+O(J_2^2), \tag{131}$$

$$\Omega' = \Omega_0' - J_2(R/p)^2\cos i\left(\frac{3}{2}u+2e\sin v+2e\sin\omega\right)+O(J_2^2), \tag{132}$$

and

$$\omega' = \omega_0' + 3J_2(R/p)^2\left[\left(1-\frac{5}{4}f\right)u+\frac{1}{24}f\sin 2u+\left(\frac{2}{3}-\frac{2}{3}f\right)(e\sin v+e\sin\omega)+\right.$$
$$\left.+\frac{1}{6}fe\sin(u+\omega)-\frac{1}{6}fe\sin\omega+\left(\frac{1}{8}f-\frac{1}{12}\right)(e^2\sin 2v+e^2\sin 2\omega)\right]+O(J_2^2). \tag{133}$$

By comparing these with the expressions for δi, $\delta\Omega$, $\delta\omega$ in Section 6 it will be seen that the amplitudes of the oscillations have been considerably reduced (in particular, the e^{-1} term in ω' is absent) and the resultant expressions are much simpler to compute.

A summary of the formulae connecting smoothed elements is given later.

13. Second-order (J_2^2) secular and long-period changes in the smoothed elements

The second-order secular and long-period changes in i' and Ω' are easily determined.

First, since

$$\begin{aligned} i' &= i+J_2\gamma_i, \\ \Delta i' &= \Delta i+J_2\Delta\gamma_i. \end{aligned} \tag{133}$$

Now setting u successively zero and 2π in (126) and subtracting, remembering that Δe, Δi are $O(J_2^2)$ but $\Delta\omega$ is $O(J_2)$,

$$\Delta\gamma_i = 3(R/p)^2\sin i\cos i\left(\frac{2}{3}e\sin\omega\right)\Delta\omega+O(J_2^2). \tag{134}$$

Thus, substituting the known values of Δi and $\Delta\omega$,

$$\Delta i' = 2\pi J_2^2(R/p)^4\left(\frac{3}{2}\sin 2i\right)\left[\left(1-\frac{5}{4}f\right)e\sin\omega+\left(\frac{7}{16}-\frac{15}{32}f\right)e^2\sin 2\omega\right]+O(J_2^2), \tag{135}$$

which differs from Δi only in the sign of the $e\sin\omega$-term.

In this expression, of course, the elements can equally well be all "dashed", since the difference between dashed and osculating elements is of order J_2.

From (87) and (88)

$$\gamma_\Omega\sin i = \gamma_i\tan u-\frac{1}{2}J_2\gamma_i^2\tan u\cot i(2+\tan^2 u)+O(J_2^2),$$

so that, since γ_i is bounded, γ_Ω is a multiple of $\sin u$, $\Delta\gamma_\Omega = 0+O(J_2^2)$ and therefore

$$\Delta\Omega' = \Delta\Omega+O(J_2^3) \tag{136}$$

and it only remains to transform the elements occurring in the expression (57) for $\Delta\Omega$ from osculating to dashed elements.

From (99), remembering that $p' = \text{constant}+O(J_2^2)$,

$$p_0 = p'(1-J_2c_p/p)$$

and

$$i_0 = i_0'-J_2c_{i0},$$

so that

$$-2\pi J_2(R/p_0)^2 \cdot \frac{3}{2}\cos i_0 = -2\pi J_2(R/p')^2\left(\frac{3}{2}\cos i_0'\right)(1+2J_2p^{-1}c_p+J_2c_{i0}\tan i)+ \\ +O(J_2^3). \tag{137}$$

Substituting the known values of c_p and c_{i0} and adding the second-order term in $\Delta\Omega$, then gives

$$\Delta\Omega' = -2\pi J_2(R/p')^2\left(\frac{3}{2}\cos i_0'\right)+ \\ +2\pi J_2^2(R/p)^4\left(\frac{3}{2}\cos i\right)\left[\left(\frac{11}{4}-\frac{7}{2}f\right)+(4-4f)e\cos\omega+\right. \\ \left.+\left(\frac{7}{4}-\frac{53}{16}f\right)e^2+\left(-\frac{7}{8}+\frac{15}{8}f\right)e^2\cos 2\omega\right]+O(J_2^3). \tag{138}$$

To determine the secular and long-period changes in the elements, ω', e', p' and σ' is considerably more laborious, although the principle is much the same as in the case of the first-order changes, leading to two equations corresponding to (100) and (101).

By variation of equations (82) and (85) it is found that

$$\frac{(1-e_0^2)^{\frac{1}{2}}}{(1+e_0\cos\omega_0)^2}\Delta\omega-\frac{(1-e_0'^2)^{\frac{1}{2}}}{(1+e_0'\cos\omega_0'^2)}\Delta\omega'-(1-e_0^2)^{\frac{1}{2}}\sin\omega_0\frac{(2+e\cos\omega_0)}{(1+e_0\cos\omega_0)^2}\Delta e+ \\ +(1-e_0'^2)^{\frac{1}{2}}\sin\omega_0'\frac{(2+e_0'\cos\omega_0')}{(1+e_0'\cos\omega_0')^2}\Delta e' \\ =\Delta\sigma'-\Delta\sigma^*-\frac{1}{2}\Delta\phi+\frac{1}{6}\frac{na^2}{\mu^2}\int_0^{2\pi}r^2f_1^2\frac{dt}{du}du+O(J_2^3). \tag{139}$$

Letting

$$\Delta\omega' = 2\pi J_2(R/p')^2 3\left(1-\frac{5}{4}f_0'\right)+2\pi J_2^2(R/p)^4\omega_{22}'. \tag{140}$$

evaluating the integral, and adding together and simplifying the known quantities, equation (139) finally gives

$$2\pi J_2^2(R/p)^4(1-e^2)\omega_{22}'-\sin\omega(2+e\cos\omega)\Delta e'+(1-e^2)^{-\frac{1}{2}}(1+e\cos\omega)^2\Delta\sigma' \\ =2\pi J_2^2(R/p)^4(1-e^2)\cdot\frac{3}{4}\left[\left(-\frac{17}{2}+\frac{77}{4}f-\frac{25}{2}f^2\right)+\left(-9+\frac{29}{2}f-\frac{45}{8}f^2\right)e\cos\omega+\right. \\ \left.+\left(-10+\frac{145}{4}f-\frac{915}{32}f^2\right)e^2+\left(7-\frac{57}{2}f+\frac{365}{16}f^2\right)e^2\cos^2\omega\right]+O(J_2^3). \tag{141}$$

A second equation, connecting $\Delta\omega'$, $\Delta e'$ and $\Delta p'$ is obtained by varying equations (84). This procedure yields the intermediate equation

$$p_0e_0'\sin\omega_0'\Delta\omega'-p_0'e_0\sin\omega_0\Delta\omega-p_0\cos\omega_0'\Delta e'+p_0'\cos\omega_0\Delta e+(1+e_0\cos\omega_0)\Delta p'- \\ -(1+e_0'\cos\omega_0')\Delta p-p_0J_2\Delta z \\ =O(J_2^3). \tag{142}$$

After determining Δz from (122), replacing $\Delta\omega'$ by the expression (140) and summing the known quantities, equation (142) gives

$$2\pi J_2^2(R/p)^4 pe\sin\omega\,.\,\omega_{22}' - p\cos\omega\Delta e' + (1+e\cos\omega)\Delta p'$$

$$= 2\pi J_2^2(R/p)^4 pe\sin\omega\,.\,\frac{3}{4}\left[\left(-8+\frac{41}{2}f-\frac{235}{16}f^2\right)+(-8+17f-10f^2)e\cos\omega+\right.$$

$$\left.+\left(-10+\frac{145}{4}f-\frac{915}{32}f^2\right)e^2+\left(\frac{15}{2}-\frac{109}{4}f+\frac{165}{8}f^2\right)e^2\cos^2\omega\right]+O(J_2^3). \quad (143)$$

Elimination of ω_{22}' between (141) and (143) gives

$$p(1+e\cos\omega)(\cos\omega-2e+e\cos^2\omega)\Delta e' - (1-e^2)(1+e\cos\omega)\Delta p' +$$

$$+pe\sin\omega(1-e^2)^{-\frac{1}{2}}(1+e\cos\omega)^2\Delta\sigma'$$

$$= 2\pi J_2^2(R/p)^4 pe\sin\omega(1-e^2)\,.\,\frac{3}{4}\left[\left(-\frac{1}{2}-\frac{5}{4}f+\frac{35}{16}f^2\right)+\right.$$

$$\left.+\left(-1-\frac{5}{2}f+\frac{35}{8}f^2\right)e\cos\omega+\left(-\frac{1}{2}-\frac{5}{4}f+\frac{35}{16}f^2\right)e^2\cos^2\omega\right]+O(J_2^3).$$

The right-hand side, not unexpectedly, is divisible by $(1+e\cos\omega)$ so that

$$p(\cos\omega-2e+e\cos^2\omega)\Delta e' - (1-e^2)\Delta p' + pe\sin\omega(1-e^2)^{-\frac{1}{2}}(1+e\cos\omega)\Delta\sigma'$$

$$= 2\pi J_2^2(R/p)^4 pe\sin\omega(1-e^2)(1+e\cos\omega)\left(-\frac{3}{8}-\frac{15}{16}f+\frac{105}{64}f^2\right)+O(J_2^3). \quad (144)$$

This equation shows that it is not possible to have all three of e', p' and σ' constant, as in the first-order approximation. The most satisfactory solution is to take e' and p' constant, so that the equations $p' = a'(1-e'^2)$ and $n'^2a'^3 = \mu$ can be maintained, and with these conditions

$$\Delta e' = \Delta p' = 0+O(J_2^3),$$

and

$$\Delta\sigma' = 2\pi J_2^2(R/p)^4(1-e^2)^{\frac{1}{2}}\,.\,\frac{3}{8}\left(-1-\frac{5}{2}f+\frac{35}{8}f^2\right)+O(J_2^3). \quad (145)$$

(140) and (143) then give

$$\Delta\omega' = 2\pi J_2(R/p')^2\,.\,3\left(1-\frac{5}{4}f_0'\right)+2\pi J_2^2(R/p)^4\,.\,6\left[\left(-1+\frac{41}{16}f-\frac{235}{8}f^2\right)+\right.$$

$$+\left(-1+\frac{17}{8}f-\frac{5}{4}f^2\right)e\cos\omega+\left(-\frac{5}{4}+\frac{145}{32}f-\frac{915}{256}f^2\right)e^2+ \quad (146)$$

$$\left.+\left(\frac{15}{16}-\frac{109}{32}f+\frac{165}{64}f^2\right)e^2\cos^2\omega\right]+O(J_2^3).$$

14. Nodal period T_Ω in terms of smoothed elements

The nodal period T_Ω is obtained by variation of equations (85) and (82). From (85),

$$\Delta E' - e'\cos E_0'(\Delta E'-2\pi)+\frac{1}{2}e'\sin E_0'(\Delta E'-2\pi)^2 = n'T_\Omega+\Delta\sigma'+O(J_2^3), \quad (147)$$

and from (82), using (83),

$$(\Delta E' - 2\pi) = -\frac{(1-e'^2)^{\frac{1}{2}}\Delta\omega'}{(1+e'\cos\omega_0')} - \frac{\frac{1}{2}(1-e'^2)^{\frac{1}{2}}e'\sin\omega_0'\Delta\omega'^2}{(1+e'\cos\omega_0')^2} + O(J_2^3). \quad (148)$$

Combining these, and using (82), (83)

$$n'T_\Omega = 2\pi - \frac{(1-e'^2)^{\frac{1}{2}}\Delta\omega'}{(1+e'\cos\omega_0')^2} - \frac{(1-e'^2)^{\frac{1}{2}}e'\sin\omega_0'\Delta\omega'^2}{(1+e'\cos\omega_0')^3} - \Delta\sigma' + O(J_2^3), \quad (149)$$

where $\Delta\omega'$, $\Delta\sigma'$ are given by (145) and (146).

At some inclination close to 63°·4 the effect of the perturbing force on T_Ω is nil, but for small inclinations the effect is relatively large, approaching 0·3 per cent.

15. Summary of formulae connecting smoothed elements

Since the practical use of smoothed elements does not involve any recourse to osculating elements it is convenient, in this summary of formulae, to drop the dashes from the smoothed elements and other variables. They now form a comprehensive system in their own right.

The motion of a satellite in the potential field

$$U = \frac{\mu}{r}\left[1 - \frac{1}{2}J_2(R/r)^2(3\cos^2\theta - 1)\right]$$

can be represented by the following formulae:

$$\left.\begin{aligned} \sin\theta\cos(\lambda-\Omega) &= \cos u \\ \sin\theta\sin(\lambda-\Omega) &= \cos i\sin u \\ \cos\theta &= \sin i\sin u \end{aligned}\right\} \quad (150)$$

$$\left.\begin{aligned} \cos E &= (e+\cos v)/(1+e\cos v) \\ \sin E &= (1-e^2)^{\frac{1}{2}}\sin v/(1+e\cos v) \\ p/r &= 1+e\cos v+J_2 z \\ E - e\sin E &= nt+\sigma, \end{aligned}\right\} \quad (151)$$

where r, θ, λ are the polar co-ordinates of the satellite, θ being geocentric co-latitude and λ right ascension;

$$\left.\begin{aligned} v &= u-\omega, \\ p &= a(1-e^2), \\ n^2a^3 &= \mu, \end{aligned}\right\} \quad (152)$$

$$a, p, e, n \text{ are constant} + O(J_2^3),$$

$$\sigma \text{ is constant} + O(J_2^2),$$

and

$$z = (R/r)^2\left[-\frac{1}{4}f\cos 2u + \left(\frac{1}{2}-\frac{3}{4}f\right)(1-e\cos v)\right] + O(J_2), \quad (153)$$

where

$$f = \sin^2 i;$$

$$i = i_0 + \frac{1}{2}J_2(R/p)^2 \sin 2i(e\cos\omega - e\cos v) +$$
$$+ \frac{3}{2}J_2^2(R/p)^4 u \sin 2i\left[\left(1-\frac{5}{4}f\right)e\sin\omega + \left(\frac{7}{16}-\frac{15}{32}f\right)e^2\sin 2\omega\right] +$$
$$+ J_2^2 \times \text{periodic terms}, \tag{154}$$

$$\Omega = \Omega_0 - J_2(R/p)^2\cos i_0\left(\frac{3}{2}u + 2e\sin v + 2e\sin\omega\right) +$$
$$+ \frac{3}{2}J_2^2(R/p)^4 u\cos i\left[\left(\frac{11}{4}-\frac{7}{2}f\right) + 4e\cos^2 i\cos\omega + \left(\frac{7}{4}-\frac{53}{16}f\right)e^2 +\right.$$
$$\left. + \left(-\frac{7}{8}+\frac{15}{8}f\right)e^2\cos 2\omega\right] + J_2^2 \times \text{periodic terms}, \tag{155}$$

$$\omega = \omega_0 + J_2(R/p)^2\left[3\left(1-\frac{5}{4}f_0\right)u + \frac{1}{8}f\sin 2u + \frac{1}{2}fe\sin(2u-v) -\right.$$
$$\left. - \frac{1}{2}fe\sin\omega + 2\cos^2 i(e\sin v + e\sin\omega) + \left(\frac{3}{8}f-\frac{1}{4}\right)(e^2\sin 2v + e^2\sin 2\omega)\right] +$$
$$+ 6J_2^2(R/p)^4 u\left[\left(-1+\frac{41}{16}f-\frac{235}{128}f^2\right) + \left(-1+\frac{17}{8}f-\frac{5}{4}f^2\right)e\cos\omega +\right.$$
$$\left. + \left(-\frac{25}{32}+\frac{181}{64}f-\frac{585}{256}f^2\right)e^2 + \left(\frac{15}{32}-\frac{109}{64}f+\frac{165}{128}f^2\right)e^2\cos 2\omega\right] +$$
$$+ J_2^2 \times \text{periodic terms}, \tag{156}$$

$$\sigma = \sigma_0 - \frac{3}{8}J_2^2(R/p)^4 u(1-e^2)^{\frac{1}{2}}\left(1+\frac{5}{2}f-\frac{35}{8}f^2\right) + J_2^2 \times \text{periodic terms}, \tag{157}$$

the zero suffices indicating conditions at the ascending node ($t = 0$, $u = 0$).

To complete the picture as far as the smoothed elements are concerned it is necessary to include the secular effects of the J_3, J_4, J_5 and J_6 terms in the potential. The elements a and n still remain constant and the changes in one revolution in e, i, Ω, ω and p due to the J_3–J_6 terms are the same for smoothed elements as for osculating elements and are therefore given by the appropriate formulae in (55)–(58) and (61).

To determine the secular change in σ, we have

$$\sigma - \sigma^* = -\int_0^t 3na\frac{F}{\mu}\,dt + O(F^2)$$

so that

$$\Delta\sigma = \Delta\sigma^* - \int_0^{2\pi} 3na\frac{F}{\mu}\cdot\frac{dt}{du}\cdot du + O(F^2)$$

or setting

$$\Delta\sigma = 2\pi[\sum_n J_n(R/p)^n\sigma_n + J_2^2(R/p)^4\sigma_{22}] \tag{158}$$

and using (59), (71),

$$\sigma_n = \sigma_n^* + 3\phi_n/2(n+1).$$

Substitution from (59) and (60) then gives

$$\sigma_2 = 0,$$

$$\sigma_3 = -\frac{3}{2}(1-e^2)^{\frac{1}{2}}\left(1-\frac{5}{4}f\right)e^{-1}\sin\omega\sin i,$$

$$\sigma_4 = \frac{9}{8}(1-e^2)^{\frac{1}{2}}\left[\left(1-5f+\frac{35}{8}f^2\right)+\frac{5}{2}\left(1-\frac{7}{6}f\right)f\cos 2\omega\right],$$

$$\sigma_5 = \frac{15}{4}(1-e^2)^{\frac{1}{2}}e^{-1}\sin i\left[\left(1-\frac{7}{2}f+\frac{21}{8}f^2\right)\left(1+\frac{9}{4}e^2\right)\sin\omega+\frac{7}{8}\left(1-\frac{9}{8}f\right)fe^2\sin 3\omega\right],$$

$$\sigma_6 = -\frac{25}{8}(1-e^2)^{\frac{1}{2}}\left[\left(1-\frac{21}{2}f+\frac{189}{8}f^2-\frac{231}{16}f^3\right)\left(1+\frac{3}{4}f^2\right)+\right.$$
$$\left.+\frac{21}{4}\left(1-3f+\frac{33}{16}f^2\right)(1-e^2)f\cos 2\omega+\frac{63}{64}\left(1-\frac{11}{10}f\right)f^2e^2\cos 4\omega\right]$$

and from (157),

$$\sigma_{22} = -\frac{3}{8}(1-e^2)^{\frac{1}{2}}\left(1+\frac{5}{2}f-\frac{35}{8}f^2\right).$$

16. Comments

There has been much controversy over the way in which the "oblateness problem" of satellite motion should be handled, the difficulties not being entirely confined to the second-order terms. King-Hele's early treatment (1958) was adequate for the purposes for which it was used but gives singularities of order J_2e in the short-period motion as a result of keeping the orbital inclination constant. The theory of Musen (1959), which has been used in the Vanguard project, is not suitable for eccentricities less than about 0·05 or greater than about 0·6. The long-period perturbations given by Brouwer (1959) and Kozai (1959) have singularities at the critical inclination ($\sin^2 i = 4/5$) so that their method cannot satisfactorily be applied, for example, to the first three Russian satellites.

The use of osculating elements is advocated by several authors, including Groves (1960) and Zhongolovitch (1960). As has already been pointed out, the short-period perturbations in the osculating elements, equations (35)–(40), besides being cumbersome are unnecessarily large in amplitude, and $\delta\omega$ has a singularity when $e = 0$, this singularity being reflected also in the J_2^2 term of the long-period perturbation. This singularity, however, is not fundamental and does not imply that the actual perigee point rotates rapidly. This is brought out in the formulae for smoothed elements, (150)–(157), developed in the present paper, which possess *no singularities at all in the J_2 and J_2^2 terms* and can therefore be used for all values of eccentricity and inclination.

There still remain singularities due to the odd-harmonics $J_3, J_5, \ldots, e^{-1}$ appearing in the secular perturbations of ω and σ^* and $\cot i$ in the secular perturbation of Ω. But these singularities are fundamental, and not just a result of the analytical method adopted. For example, when i is small (near-equatorial orbit) the odd harmonics in the gravitational field cause a rapid movement of the ascending node.

Acknowledgments

The author wishes to thank his colleagues Dr R. N. A. Plimmer, who independently calculated certain of the secular changes in the osculating elements, and Mr D. G. King-Hele for his valuable comments and criticisms.

This paper is published by permission of the Controller, H.M. Stationery Office, and Crown Copyright is reserved.

Royal Aircraft Establishment,
Farnborough,
Hants:
1960, *July* 28.

References

Brenner, J. L., 1959. *Stanford Res. Inst. Report ASTIA*, No. AD-216455.
Cook, A. H., 1958. *Geophys. J.*, **3**, 341.
Groves, G. V., 1960. *Proc. Roy. Soc. A*, **254**, 48.
Jeffreys, H., 1952. *The Earth.* (3rd ed.) (Cambridge: The University Press.)
King-Hele, D. G., 1958. *Proc. Roy. Soc. A*, **247**, 49.
King-Hele, D. G., 1960. *Nature* (*London*), **187**, 490.
King-Hele, D. G. & Walker, D. M. C., 1960. *Nature* (*London*), **186**, 928.
Merson, R. H. & King-Hele, D. G., 1958. *Nature* (*London*), **182**, 640.
Musen, P., 1959. *J. Geophys. Res.*, **64**, 2271.
Smart, W. M., 1953. *Celestial Mechanics.* (Longmans, Green & Co. Ltd.)
Zhongolovitch, I. D., 1960. *Bull. Inst. Theor. Astr.* (Leningrad), VII, **7**, 521.

Resonant Orbits of Artificial Satellites and Longitude Terms in the Earth's External Gravitational Potential*

A. H. Cook

(with an Appendix by H. J. Norton)
Communication from the National Physical Laboratory

"Every arrow that flies feels the attraction of the Earth."—*Longfellow, Hiawatha.*

Summary

The zonal harmonics in the Earth's external gravitational potential give rise to long period or secular perturbations of the orbits of close artificial satellites and can be estimated from such perturbations but the tesseral and sectorial harmonics in general produce only short period perturbations which are far more difficult to observe. There is thus little prospect at present of finding the longitude-dependent parts of the potential from the motion of an arbitrary satellite and so it is of interest to see if in special circumstances long period or secular perturbations could arise from these parts. This paper contains a preliminary study of orbits with periods bearing some specific relation to the period of the Earth's rotation. For the purposes of this study, it has been assumed that the eccentricity of the orbit and its inclination to the equator are so small that only the first power of the eccentricity and the square of the inclination need be retained and it is also supposed that, independently of the longitude terms in the potential, the longitudes of the node and perigee change linearly with time.

It is found that secular and long period perturbations can arise and that they differ according to the parity of the difference $(p-q)$ for the associated Legendre function $P_q^p(\cos\theta)$ in the spherical harmonic expansion of the potential. One approximate condition for such perturbations is that

$$(q-1)n = q\omega_r$$

where n is the mean motion of the satellite and ω_r the Earth's spin angular velocity. The condition is slightly different for odd and even $(p-q)$.

When $(p-q)$ is even, there are perturbations of order e^{-1} (e is the eccentricity of the orbit) in the mean anomaly and the longitude of perigee while when $(p-q)$ is odd there are perturbations of order

* A preliminary version of this paper was presented to the 1st International Space Science Symposium (Cook 1960).

$(\sin i)^{-1}$ (i is the inclination of the orbit to the equator) in the longitude of the node and of perigee.

The simple case of the ellipticity of the equator ($P_2^2(\cos\theta)$) is discussed in rather more detail and some preliminary consideration is given to the problems of the realization of resonant orbits.

The general expressions for the perturbations, excluding terms proportional to e^2 and $\sin^4 i$ or of higher order, are given in the Appendix where the complete set of resonance conditions is listed. The resonances discussed in the body of the paper are generally the most important in that they occur for the smallest orbits but there are resonances with smaller orbits in some special cases.

1. Introduction

Although the zonal harmonics in the Earth's external gravitational potential can be found from observations of the secular and long period perturbations of the orbits of close artificial satellites, the parts of the potential that vary with longitude cannot be so found because in general, as a result of the rotation of the Earth, they give rise only to short period perturbations and not to long period or secular ones. Short period perturbations, having periods equal to those of the rotation of the satellite in its orbit or of the Earth about its axis, or combinations thereof, are difficult to observe, and cannot be found from observations at a single station. It is therefore of interest to see if long period or secular perturbations can arise in special circumstances and one naturally thinks of the possibility of resonant orbits where the period is in some definite relation to the period of the Earth's rotation about its axis. In this paper I attempt a preliminary study of such orbits; I consider only the theoretical problem of the origin of secular or long period terms and do not discuss the technical problems of putting satellites into such orbits. Resonant orbits have also been discussed by Groves (1960).

2. Notation and method

I suppose that the orbit can be defined in terms of osculating elements at the ascending node and take these to be:

a semi-major axis
e eccentricity
τ time of perigee passage
ω longitude of perigee measured from ascending node
Ω longitude of ascending node referred to axes fixed in space
i inclination of orbit to the equator.

n is the mean motion of the satellite
M the mean anomaly is $n(t-\tau)$ or $nt+\chi$
v is the true anomaly
r,θ,λ are spherical polar co-ordinates referred to axes fixed in direction in space with the centre at the centre of mass of the Earth and with the polar axis the Earth's axis of rotation
a_e is the Earth's equatorial radius
ω_r is the Earth's spin angular velocity.

Then I take a typical term of the perturbing potential to be

$$V_{pq} = J_{pq}\frac{a_e^p}{r^{p+1}}P_p^q(\cos\theta)\sin q(l+\beta_{pq})$$

where l is the longitude measured from axes fixed in the Earth. J_{pq} and β_{pq} are constants and P_p^q is an associated Legendre function.

Lagrange's equations for the variations of the osculating elements are

$$\dot{a} = \frac{2}{na}\frac{\partial V_{pq}}{\partial \chi}; \tag{1(a)}$$

$$\dot{e} = \frac{1}{na^2e}\left\{(1-e^2)\frac{\partial V_{pq}}{\partial \chi} - (1-e^2)^{\frac{1}{2}}\frac{\partial V_{pq}}{\partial \omega}\right\}; \tag{1(b)}$$

$$\dot{\chi} = -\frac{1-e^2}{na^2e}\frac{\partial V_{pq}}{\partial e} - \frac{2}{na}\frac{\partial V_{pq}}{\partial a}; \tag{1(c)}$$

$$\dot{\Omega} = \frac{1}{na^2(1-e^2)^{\frac{1}{2}}}\frac{1}{\sin i}\frac{\partial V_{pq}}{\partial i}; \tag{1(d)}$$

$$\dot{\omega} = \frac{(1-e^2)^{\frac{1}{2}}}{na^2e}\frac{\partial V_{pq}}{\partial e} - \frac{\cot i}{na^2(1-e^2)^{\frac{1}{2}}}\frac{\partial V_{pq}}{\partial i}; \tag{1(e)}$$

$$\frac{di}{dt} = \frac{1}{na^2(1-e^2)^{\frac{1}{2}}}\left(\cot i\frac{\partial V_{pq}}{\partial \omega} - \operatorname{cosec} i\frac{\partial V_{pq}}{\partial \Omega}\right). \tag{1(f)}$$

In order to differentiate V_{pq} with respect to the orbital elements, expressions for r, $\cos\theta$ and λ in terms of these elements are required. We have

$$\left(\frac{a}{r}\right)^p = \{1+2\sum_{\nu=1}^{\infty}J_\nu(\nu e)\cos\nu M\}^p, \qquad \text{(Smart 1953, p. 41)}$$

where J_ν is the Bessel function of order ν. Also

$$\cos\theta = \sin(\omega+v)\sin i.$$

To determine l take the zero longitude in the system of fixed axes to coincide with the longitude of Greenwich at $t = 0$. The longitude λ_G of Greenwich, relative to fixed axes at any time t is then $\omega_\gamma t$.

Let the longitude of the satellite measured along the equator from the node be ψ. Then λ, the longitude of the satellite measured along the equator from the origin of fixed axes, is

$$\psi+\Omega.$$

But

$$\tan\psi = \tan(\omega+v)\cos i$$

and for small inclinations

$$\psi = \omega+v-\tfrac{1}{4}\gamma^2\sin 2(\omega+v)+\ldots$$

where $\gamma = \tan i$ (Smart 1953, p. 101). Hence

$$l = \lambda-\lambda_G = \omega+\Omega+v-\omega_\gamma t-\tfrac{1}{4}\gamma^2\sin 2(\omega+v)+\ldots.$$

Then

$$\frac{\partial}{\partial e}\left(\frac{a}{r}\right)^p = 2p\{1+2\sum J_\nu(\nu e)\cos\nu M\}^{p-1}\sum \nu J_\nu'(\nu e)\cos\nu M,$$

$$\frac{\partial}{\partial \chi}\left(\frac{a}{r}\right)^p = \frac{\partial}{\partial M}\left(\frac{a}{r}\right)^p = -2p\{1+2\sum J_\nu(\nu e)\cos\nu M\}^{p-1}\sum \nu J_\nu(\nu e)\sin\nu M,$$

and

$$\left(\frac{\partial}{\partial \omega}, \frac{\partial}{\partial \Omega}, \frac{\partial}{\partial i}, \frac{\partial}{\partial a}\right)\left(\frac{a}{r}\right)^p = 0.$$

Also,

$$\frac{\partial}{\partial \omega}(\cos\theta) = \cos(\omega+v)\sin i,$$

$$\frac{\partial}{\partial i}(\cos\theta) = \sin(\omega+v)\cos i,$$

$$\frac{\partial}{\partial \chi}(\cos\theta) = \cos(\omega+v)\sin i\frac{\partial v}{\partial M},$$

$$\frac{\partial}{\partial e}(\cos\theta) = \cos(\omega+v)\sin i\frac{\partial v}{\partial e},$$

and

$$\left(\frac{\partial}{\partial a}, \frac{\partial}{\partial \Omega}\right)\cos\theta = 0.$$

Since

$$v = M+(2e-\tfrac{1}{4}e^3)\sin M+\left(5e^2-\frac{11}{24}e^4\right)\sin 2M+\ldots, \qquad \text{(Smart 1953, p. 38)}$$

then

$$\frac{\partial v}{\partial M} = 1+(2e-\tfrac{1}{4}e^3)\cos M+2\left(5e^2-\frac{11}{24}e^4\right)\cos 2M+\ldots,$$

and

$$\frac{\partial v}{\partial e} = (2-\tfrac{3}{4}e^2)\sin M+\left(10e-\frac{11}{6}e^3\right)\sin 2M+\ldots.$$

Lastly,

$$\frac{\partial l}{\partial \Omega} = 1,$$

$$\frac{\partial l}{\partial \omega} = 1-\tfrac{1}{2}\gamma^2\cos 2(\omega+v),$$

$$\frac{\partial l}{\partial \chi} = \{1-\tfrac{1}{2}\gamma^2\cos 2(\omega+v)\}\frac{\partial v}{\partial M},$$

$$\frac{\partial l}{\partial e} = \{1-\tfrac{1}{2}\gamma^2\cos 2(\omega+v)\}\frac{\partial v}{\partial e},$$

$$\frac{\partial l}{\partial i} = -\tfrac{1}{2}\tan i\sec^2 i\sin 2(\omega+v)$$

and

$$\frac{\partial l}{\partial a} = 0.$$

I shall consider only orbits with small inclinations and eccentricities, namely those for which $\tan^2 i$ and e are less than 0·1 and all expressions will therefore be carried only so far as the first order in $\tan^2 i$ and e, except where they are to be divided by e or $\sin i$.

3. Ellipticity of the Equator

The potential V_{22} corresponds to an elliptical equator. It is of particular geodetic and geophysical interest and so is a useful simple case to discuss first.

The procedure that will be followed is to pick out from the differentials of V_{22} the terms of long period and to see how they combine in the expressions for the rates of change of the orbital elements.

Thus,

$$\frac{\partial V_{22}}{\partial i} = J_{22}\frac{a_e^2}{r^3}\begin{bmatrix} -6\sin^2(\omega+v)\sin i\cos i\sin 2(l+\beta_{22}) \\ -3\tan i\sec^2 i\sin 2(\omega+v)\cos 2(l+\beta_{22}) \\ +3\sin^2(\omega+v)\sin 2(\omega+v)\sin^2 i\tan i\sec^2 i\cos 2(l+\beta_{22}) \end{bmatrix}.$$

The factor a_e^2/r^3 is equal to

$$\frac{1}{a_e}\left(\frac{a_e}{a}\right)^3\left(\frac{a}{r}\right)^3$$

and $(a/r)^3$ is equal to $1+3e\cos M+\ldots$. Thus $\partial V_{22}/\partial i$ contains terms of order $\sin i$, $\sin^3 i$ and $e\sin i$, together with terms of higher order. Taking these successively, and omitting the common factors $(J_{22}/ae)[(ae/a)^3]$, we have:

terms of order sin i:

$$-6\sin^2(\omega+v)\sin i\cos i\sin 2(l+\beta_{22}).$$

This has a secular part if $2(\omega+v)$ is equal to $2l$, that is, if

$$\omega+v = \omega+v+\Omega-\omega_r t-\tfrac{1}{4}\gamma^2\sin 2(\omega+v)+\ldots.$$

This condition cannot in fact be satisfied because $\dot{\Omega}$ is only about 4 deg/day for close satellites and so this term has no long period or secular component. The same argument applies to the term

$$-3\tan i\sec^2 i\sin 2(\omega+v)\cos 2(l+\beta_{22})$$

and to the term of order $\sin^3 i$, namely

$$3\sin^2 i\tan i\sec^2 i\sin^2(\omega+v)\sin 2(\omega+v)\cos 2(l+\beta_{22}).$$

The terms of order e sin i are

$$-18e\sin i\cos i\sin^2(\omega+v)\sin 2(l+\beta_{22})\cos M$$

and

$$-9e\sin i\sec^3 i\sin 2(\omega+v)\cos 2(l+\beta_{22})\cos M.$$

The former gives rise to a secular term if

$$M = 2(l+\beta_{22}).$$

In general this will not be satisfied exactly; accordingly let

$$2(l+\beta_{22}) = M+\eta,$$

when the term may be written as

$$-9e\sin i\cos i\{1-\cos 2(\omega+v)\}(\sin M\cos\eta+\cos M\sin\eta)\cos M.$$

Then the long period part is

$$-\frac{9}{2}e\sin i\cos i\sin\eta.$$

The second term gives rise to a long period component if

$$2(l+\beta_{22}) = 2(\omega+v)-M+\eta'.$$

I shall confine myself in this paper to terms containing $\sin\eta$ or $\cos\eta$ since in general these are the largest terms, and also the condition $M = 2(l+\beta_{22})$ corresponds to an orbit of smaller radius than for other conditions. Dr Norton gives the complete expansions in his Appendix.

Proceeding in this way, the terms proportional to $\cos\eta$ or $\sin\eta$ in all the differentials are identified and listed in Table 1. Now from the equation for $\dot{\Omega}$, it follows that the part of $\dot{\Omega}$ proportional to $\sin\eta$ is

$$-\frac{9}{2}J_{22}\frac{e}{na^3(1-e^2)^{\frac{1}{2}}}\left(\frac{a_e}{a}\right)^2\cos i\sin\eta.$$

Similarly, all the parts of the rates of change of the orbital elements proportional to $\sin\eta$ or $\cos\eta$ are picked out and listed in Table 2.

Table 1

Factors of terms proportional to sin η or cos η in derivatives of V_{22}
$\eta = 2(l+\beta_{22})-M$

Derivative with respect to	Order of term: 1	e	γ^2	γe
a	—	$-\frac{27}{2a}e\sin\eta$	—	
e	$-\frac{3}{2}\sin\eta$	—	$\frac{3}{4}\sin^2 i\sin\eta$	
χ	—	$\frac{21}{2}e\cos\eta$	—	
Ω	—	$9e\cos\eta$	—	
ω	—	$9e\cos\eta$	—	
i	—	—	—	$\frac{9}{2}e\sin i\cos i\sin\eta$

All terms to be multiplied by $\frac{J_{22}}{a_e}\left(\frac{a_e}{a}\right)^3$

Table 2

Factors of terms proportional to sin η or cos η in rates of change of elements. $\eta = 2(l+\beta_{22})-M$

Element	Order of term e^{-1}	1	e	$\gamma^2 e^{-1}$	γe
$\dot{a}$	—	—	$\frac{21}{na} e \cos\eta$	—	—
$\dot{e}$	—	$\frac{3}{2na^2}\cos\eta$	—	—	—
$\dot{\chi}$	$\frac{3}{2na^2}\frac{\sin\eta}{e}$	—	$\frac{27}{na^2} e \sin\eta$	$-\frac{3}{4na^2}\frac{\sin^2 i}{e}\sin\eta$	—
$\dot{\Omega}$	—	—	$-\frac{9}{2na^2} e \cos i \sin\eta$	—	—
$\dot{\omega}$	$-\frac{3}{2na^2}\frac{\sin\eta}{e}$	—	$\frac{9}{2na^2} e \cos^2 i \sin\eta$	$\frac{3}{4na^2}\frac{\sin^2 i}{e}\sin\eta$	—
$\frac{di}{dt}$	—	—	—	—	$-\frac{9}{na^2} e \tan\frac{1}{2}i \cos\eta$

All terms to be multiplied by $\frac{J_{22}}{a}\left(\frac{a_e}{a}\right)^3$

Let us now look in more detail at the condition

$$2(l+\beta_{22}) = M+\eta.$$

Since $v = M+2e\sin M+\frac{5}{4}e^2\sin 2M+\ldots$, the condition is

$$M+\eta = 2\{\omega+\Omega+M-\omega_\gamma t+2e\sin M-\tfrac{1}{4}\gamma^2\sin 2(\omega+M)+\ldots\}+2\beta_{22}.$$

Suppose, as is the case with close satellites, that ω and Ω are very nearly linear functions of time, so that

$$\omega = \omega_0+\dot{\omega}t+\text{small periodic terms}$$

and

$$\Omega = \Omega_0+\dot{\Omega}t+\text{small periodic terms.}$$

Also, since $M = nt+\chi$,

$$nt+\chi+\eta = 2\{\omega_0+\Omega_0+nt+\chi-\omega_\gamma t+2e\sin M+\dot{\omega}t+\dot{\Omega}t-\tfrac{1}{4}\gamma^2\sin 2(\omega+M)\}+ \\ +2\beta_{22}+\text{periodic terms.}$$

Thus we must have

$$n+2(\dot{\omega}+\dot{\Omega}) = 2\omega_\gamma$$

as the condition for resonance.

Then

$$\eta = 2(\omega_0+\Omega_0+\beta_{22})+\chi+4e\sin M-\tfrac{1}{4}\gamma^2\sin 2(\omega+M).$$

The small terms with argument M are periodic with speed n.

We then see that the magnitude of the secular contribution to $\dot{\Omega}$, which is proportional to $\sin\eta$, is determined by the value of χ, that is, by the position of the satellite at zero time.

In particular, $\sin\eta = 0$ if

$$\chi = j\pi-2(\omega_0+\Omega_0+\beta_{22}) \qquad (j = 0, 1,\ldots)$$

(ignoring the small terms in γ^2 and e). But the longitude λ_s of the satellite at $t = 0$ is $\omega_0+\Omega_0+\chi$ (for very small inclinations) and so

$$\lambda_s = j\pi-\omega_0-\Omega_0-2\beta_{22}.$$

Similarly, for a maximum effect, positive or negative

$$\lambda_s = (2j+1)\frac{\pi}{2}-\omega_0-\Omega_0-2\beta_{22}.$$

It follows from these considerations that if J_{22} and β_{22} are both unknown, they cannot be determined from a single orbit.

However, if η changes slowly with time because the resonance condition is not fulfilled exactly, $\sin\eta$ will pass through the values $+1$ and -1, so that J_{22} could be found from the amplitude of the periodic variation of $\dot{\Omega}$. Further, β_{22} could be found from the satellite position at the maximum.

The variation of $\dot{\Omega}$ is only of order eJ_{22} and so scarcely likely to be observed. It can be seen from Table 2 that the largest variations are in $\dot{\chi}$ and $\dot{\omega}$ and are of

order $e^{-1}J_{22}$. Terms proportional to e^{-1} are a novel feature of these resonant orbits, they arise from differentiation of $(a/r)^3$ and $\sin 2(l+\beta_{22})$ with respect to e. Since the terms in $\dot{\chi}$ and $\dot{\omega}$ are of opposite sign, the net effect on the longitude of the satellite is zero, but motion of perigee should be observable. The next largest term is that in $\dot{e}$ of order J_{22}. All other terms of are order eJ_{22} or $J_{22}\sin^2 i$.

Two limitations on the analysis must be emphasized. First, the inclination of the orbit has been restricted by assuming that $\tan^2 i$ is less than 0·1, and no attempt has been made to discover what features are characteristic of orbits with large inclinations. Secondly, it has been assumed, particularly in writing down the condition for resonance, that the variation of ω is dominated by the secular part such as arises with close satellites from the principal second harmonic in the Earth's gravitational field. But, for a satellite with a period roughly twice that of the Earth's rotation, namely at a distance of about 4 Earth's radii, the perturbations due to the Sun and Moon are important and the validity of this assumption of a dominant secular term may require study.

4. Resonant orbits for a general potential

In this section I concentrate on determining the resonance condition for a disturbing potential of general form,

$$V_{pq} = \frac{J_{pq}}{a_e}\left(\frac{a_e}{a}\right)^{p+1}\left(\frac{a}{r}\right)^{p+1} P_p^q(\cos\theta)\sin q(l+\beta_{pq}).$$

Differentiation of the factors a^{-p}, $(a/r)^p$ and $\sin q(l+\beta_{pq})$ is quite straightforward. The differentials of $(a/r)^p$ were given in Section 2 and those of $\sin q(l+\beta_{pq})$ are just

$$q\cos q(l+\beta_{pq})\left(\frac{\partial}{\partial\Omega}, \frac{\partial}{\partial\omega}, \frac{\partial}{\partial\chi}, \frac{\partial}{\partial e}, \frac{\partial}{\partial i}\right)l.$$

The associated Legendre function $P_p^q(\cos\theta)$ is

$$\frac{(\sin\theta)^q}{2^p p!}\frac{d^{p+q}}{d(\cos\theta)^{p+q}}(\cos^2\theta-1)^p$$

which may be written as

$$(\sin\theta)^q \sum_{r=0} a_r(\cos\theta)^{p-q-2r}.$$

the upper limit in the summation being $\frac{1}{2}(p-q)$ or $\frac{1}{2}(p-q-1)$ according as $(p-q)$ is even or odd. a_r is

$$\frac{(-)^r(2p-2r)!}{2^p r!(p-r)!(p-q-2r)!}.$$

If we ignore terms in $\sin^4 i$ or higher order, we may write

$$\frac{dP_p^q}{d\cos\theta} = -Xx\cos\theta \qquad \text{if } (p-q) \text{ is even}$$

and

$$= Y[1-y\cos^2\theta] \qquad \text{if } (p-q) \text{ is odd,}$$

where

$$X = \frac{(-)^{\frac{1}{2}(p-q)}(p+q)!}{2^p\{\frac{1}{2}(p-q)\}!\{\frac{1}{2}(p+q)\}!},$$

$$x = (p^2+p-q^2),$$

$$Y = \frac{(-)^{\frac{1}{2}(p-q-1)}(p+q+1)!}{2^p\{\frac{1}{2}(p-q-1)\}!\{\frac{1}{2}(p+q+1)\}!}$$

and

$$y = \tfrac{1}{2}(p^2+p-q^2-1).$$

We also ignore terms of e^2 or $\sin^4 i$ in the other factors of the differentials of V_{pq}. Thus we write

$$1+2\sum J_\nu(\nu e)\cos\nu M = 1+e\cos M$$

and

$$\left(\frac{\partial}{\partial e}, \frac{\partial}{\partial \chi}\right)\{1+2\sum J_\nu(\nu e)\cos\nu M\}^{p+1} = (p+1)(1+e\cos M)^p \times$$

$$\times(\cos M+2e\cos 2M-e\sin M).$$

With the foregoing restrictions on e and $\sin^2 i$ terms in the differentials of V_{pq} and in the rates of change of the orbital elements are given in Tables 3(a) and (b) and 4(a) and (b).

It will be seen that the behaviour differs according to whether $(p-q)$ is even or odd. When $(p-q)$ is even, there is a resonance condition analogous to that for the V_{22} term, namely

$$q(l+\beta_{pq}) = M+\eta$$

or

$$(q-1)n+q(\dot\omega+\dot\Omega) = q\omega_r,$$

with η equal to

$$q(\omega_0+\Omega_0+\beta_{pq})+(q-1)\chi+2qe\sin M-\tfrac{1}{4}q\gamma^2\sin 2(\omega+M)+\ldots.$$

But when $(p-q)$ is odd, the corresponding resonance condition is

$$q(l+\beta_{pq}) = \omega+v+\zeta,$$

that is,

$$q[\omega_0+\Omega_0+\chi+nt-\omega_r t+2e\sin M+\dot\omega t+\dot\Omega t-\tfrac{1}{4}\gamma^2\sin 2(\omega+v)+\beta_{pq}]$$
$$= \omega_0+\dot\omega t+\chi+nt+2e\sin M+\zeta.$$

Hence

$$(q-1)(n+\dot\omega) = q(\omega_r-\dot\Omega)$$

while ζ is equal to

$$(q-1)(\omega_0+\chi)+q(\Omega_0+\beta_{pq})+2(q-1)e\sin M-\tfrac{1}{4}q\gamma^2\sin 2(\omega+M).$$

The complete set of resonance conditions is given by Dr Norton in the Appendix.

In addition, a whole series of other resonances occurs. $(a/r)^{p+1}$ may be expanded in a series of terms such as $e^\nu\sin\nu M$ and in general these terms give rise to resonances when

$$\nu(M+\eta_\nu) = q(l+\beta_{pq})$$

Table 3(a)

Factors of terms proportional to sin η or cos η in derivatives of V_{pq} for $(p-q)$ even. $\eta = q(l+\beta_{pq})-M$

Derivatives with respect to	Order of terms: 1	e	γ^2	$e\gamma$
a	—	$-\frac{1}{2}\frac{(p+1)^2}{a}e\sin\eta$	—	—
e	$\frac{1}{2}(p-2q+1)\sin\eta$	—	$-\frac{1}{4}(p-2q+1)(p^2+p-q^2)\sin^2 i\sin\eta$	—
χ	—	$\frac{1}{2}(pq+3q-p-1)e\cos\eta$	—	—
Ω	—	$\frac{1}{2}q(p+1)e\cos\eta$	—	—
ω	—	$\frac{1}{2}q(p+1)e\cos\eta$	—	—
i	—	—	—	$-\frac{1}{4}(p+1)(p^2+p-q^2)e\sin i\cos i\sin\eta$

All terms to be multiplied by $\frac{J_{pq}}{a_e}\left(\frac{a_e}{a}\right)^{p+1}\frac{(-)^{\frac{1}{2}(p-q)}(p+q)!}{2^p[\frac{1}{2}(p-q)]![\frac{1}{2}(p+q)]!}$

Table 3(b)

Factors of terms proportional to sin ζ or cos ζ in derivatives of V_{pq} for $(p-q)$ odd. $\zeta = q(l+\beta_{pq})-(\omega+v)$

Derivatives with respect to	Order of terms: 1	γ	γ^2	$e\gamma$
a	—	$-\frac{1}{2a}(p+1)\sin i\cos\zeta$	—	—
e	—	—	—	$\frac{1}{4}p(p+1)e\sin i\cos\zeta$
χ	—	$-\frac{1}{4}(q-1)\sin i\sin\zeta$	—	—
Ω	—	$-\frac{1}{4}q\sin i\sin\zeta$	—	—
ω	—	$-\frac{1}{4}(q-1)\sin i\sin\zeta$	—	—
i	$\frac{1}{2}\cos i\cos\zeta$	—	$-\{\frac{3}{16}(p^2+p-q^2-2)\cos i+\frac{1}{4}q\sec i^3\}\sin^2 i\cos\zeta$	—

All terms to be multiplied by $\dfrac{J_{pq}}{a_e}\left(\dfrac{a_e}{a}\right)^{p+1}\dfrac{(-)^{\frac{1}{2}(p-q-1)}(p+q+1)!}{2^p[\frac{1}{2}(p-q-1)]![\frac{1}{2}(p+q+1)]!}$

Table 4(a)

Factors of terms proportional to sin η or cos η in rates of change of elements for (p−q) even. $\eta = q(l+\beta_{pq})-M$

Element	Order of term e^{-1}	1	e	$\gamma^2 e^{-1}$	$e\gamma$
$\dot{a}$	—	—	$\frac{1}{na}(pq+3q-p-1)e\cos\eta$	—	—
$\dot{e}$	—	$\frac{1}{2na^2}(2q-p-1)\cos\eta$	—	—	—
$\dot{\chi}$	$-\frac{1}{2na^2}(p-2q+1)\frac{\sin\eta}{e}$	—	$\frac{1}{na^2}(p+1)^2 e\sin\eta$	$+\frac{1}{8na^2}(p-2q+1)\times$ $\times(p^2+p-q^2)\frac{\sin^2 i}{e}\sin\eta$	—
$\dot{\Omega}$	—	—	$-\frac{1}{4na^2}(p+1)\times$ $\times(p^2+p-q^2)e\cos i\sin\eta$	—	—
$\dot{\omega}$	$\frac{1}{2na^2}(p-2q+1)\frac{\sin\eta}{e}$	—	$+\frac{1}{4na^2}(p+1)\times$ $\times(p^2+p-q^2)e\cos^2 i\sin\eta$	$-\frac{1}{8na^2}(p-2q+1)\times$ $\times(p^2+p-q^2)\frac{\sin^2 i}{e}\sin\eta$	—
$\frac{di}{dt}$	—	—	—	—	$-\frac{1}{2na^2}q(p+1)\tan(\frac{1}{2}i)e\cos\eta$

All terms to be multiplied by $\frac{J_{pq}}{a_e}\left(\frac{a_e}{a}\right)^{p+1}\frac{(-)^{\frac{1}{2}(p-q)}(p+q)!}{2^p[\frac{1}{2}(p-q)]![\frac{1}{2}(p+q)]!}$

Table 4(b)

Factors of terms proportional to sin ζ or cos ζ in rates of change of elements for $(p-q)$ *odd*

$\zeta = q(l+\beta_{pq})-(\omega+v)$

Element	Order of terms γ^{-1}	1	γ
a	—	—	$-\frac{1}{na}(q-1)\sin i\sin\zeta$
e	—	—	—
χ	—	—	$-\frac{1}{4na^2}(p-4)(p+1)\sin i\cos\zeta$
Ω	$\frac{1}{2na^2\gamma}\cos\zeta$	—	$-\frac{1}{16na^2}[3(p^2+p-q^2-2)\cos i+2q\sec^2 i]\sin i\cos\zeta$
ω	$-\frac{1}{2na^2}\frac{\cos i}{\gamma}\cos\zeta$	—	$\frac{1}{16na^2}[3(p^2+p-q^2-2)\cos^2 i+2q\sec^2 i+4p(p+1)]\sin i\cos\zeta$
i	—	$\frac{1}{2na^2}[q-(q-1)\cos i]\sin\zeta$	—

All terms to be multiplied by $\frac{J_{pq}}{a_e}\left(\frac{a_e}{a}\right)^{p+1}\frac{(-)^{\frac{1}{2}(p-q-1)}(p+q+1)!}{2^p[\frac{1}{2}(p-q-1)]![\frac{1}{2}(p+q+1)]!}$

and if e is close to 1, many of these terms may be important, whatever the value of i. Such resonances may therefore be important with highly eccentric orbits.

Furthermore, $P_p^q(\cos\theta)$ may be expanded in powers of $\sin(\omega+v)\sin i$ and terms of order greater than $\sin^2 i$ will give rise to resonances, but the situation is now more complex because of the dependence on $\sin^2 i$ which has also to be taken into account in the expression for l.

Thus, in general, resonances will be obtained whenever

$$\nu n = q(n-\omega_\tau) \text{ approximately,}$$

or

$$n = \frac{q\omega_\tau}{q-\nu},$$

the exact condition differing slightly for $(p-q)$ odd or even.

Since the same value of the ratio $q/(q-\nu)$ is obtained for an infinite set of pairs of q and ν, it is clear that in general an infinite set of harmonics contributes to the rates of change of elements at any one resonance. Also, when ν is 1 and q is quite large, all the resonances cluster round the value

$$n = \omega_\tau$$

and there will be a sort of band spectrum with this frequency as a point of accumulation, contributions to the rates of change of the elements again coming from an infinite set of harmonics.

The situation is much simpler when the eccentricity is small for then only the first order resonance gives terms that are not very small.

5. Realization of orbits

It is evident that if we consider only the terms in Tables 2 and 4, the orbits should have small eccentricities and inclinations so that the perturbations of order e^{-1} and cosec i may be as large as possible. However e and i cannot be too small or the effects, the motions of ω and Ω, will not be observable since it will become impossible to define perigee and the nodes observationally. The periods giving such resonances range from 12 h for $q = 2$, to 24 h. Some of these and the corresponding semi-major axes are listed in Table 5. The velocity of a satellite

Table 5

Periods, semi-major axes and velocities for $E = \pi/2$ for resonant orbits

q	T	a	u_E
	h	Earth's radii	km/s
2	12	4·2	4
3	16	5·1	4·4
4	18	5·5	4·6
5	19	5·75	4·7
10	22	6·2	4·9
∞	24	6·65	5·04

in its orbit is the most relevant quantity in considering the actual realization of an orbit and in Table 5 the velocities u_E at eccentric angles of $\pi/2$ are given since these are independent of eccentricity (this velocity is $(fM_e/a)^{\frac{1}{2}}$, M_e being the

mass of the Earth and f the gravitational constant). The important fact shown in Table 5 is the small range of velocities corresponding to the resonant orbits. This may be shown in another way. If δu_E is the departure of u_E from the resonant value, the period T_η of the long-period motion proportional to $\sin\eta$ or $\cos\eta$ is

$$\frac{2\pi}{3(q-1)}\frac{a}{\delta u_E}.$$

We then find the following values of T_η:

for $q = 2$,	$\delta u_E = 6{\cdot}6\times10^{-3}$km/s,	$T_\eta = 100d$;
	$\delta u_E = 6{\cdot}6\times10^{-2}$km/s,	$T_\eta = 10d$;
for $q = 4$,	$\delta u_E = 2{\cdot}85\times10^{-2}$km/s,	$T_\eta = 10d$;
for $q = 10$,	$\delta u_E = 1{\cdot}06\times10^{-2}$km/s,	$T_\eta = 10d$.

Clearly it will be more difficult to put a satellite into resonant orbits corresponding to the higher values of q.

From other points of view, the observation and analysis of resonant orbits may not be too difficult. Whether or not the velocity may be adjusted within such close limits so that T_η may be chosen to be quite different from other periods such as the revolution of the node or perigee or the Moon's orbital period and so on, the actual value of u_E can be measured and T_η identified and perturbations with this period can be separated out. There will be no non-resonant contributions to perturbations with this period and therefore it is not necessary to have a highly exact theory of the effect of the Earth's oblateness and of the Sun and the Moon, in order to determine longitude terms.

It would be an advantage to have an orbit of small inclination and small eccentricity, always provided that perigee and the nodes may be defined by observation.

Some recent results of Kozai & Whitney (1959) show that the effects of the Sun and Moon will have to be studied rather carefully. Kozai & Whitney found that the eccentricity and perigee distance of Satellite 1959 δ_2 (the apogee distance of which is 25 000 miles) are considerably perturbed by the Sun and Moon and although the period does not appear to be greatly affected, it may be that if a resonant orbit were subject to such large perturbations, the resonance condition would be destroyed. The perturbations computed by Kozai & Whitney are proportional to the eccentricity of the orbit and one may therefore expect that if the eccentricity is small, the perturbations due to the Sun and Moon may not disturb the resonance conditions. A numerical investigation might help to elucidate the matter.

Acknowledgments

I am indebted to Dr H. J. Norton for correcting a number of errors in the draft of this paper and for indicating some simplification.

This paper has been prepared as part of the research programme of the National Physical Laboratory and is published with the permission of the Director of the Laboratory.

National Physical Laboratory,
Teddington,
Middlesex:
1960 *September.*

References

Cook, A. H., 1960. *Proc. 1st Int. Space Sci. Symposium, Nice.* (Amsterdam: N. Holland Publ. Co.)
Groves, G. V., 1960. *Proc. Roy. Soc. A*, **254**, 48.
Kozai, Y. & Whitney, C. A., 1959. *Res. in Space Sci., Sp. Rep. No.* 30, p. 1.
Smart, W. M., 1953. *Celestial Mechanics.* (London: Longmans.)

Appendix

H. J. Norton

A complete set of resonance conditions

In this appendix it is shown that many resonant orbits are possible besides those given in Section 4 of the paper. In general they will either be further from the Earth or else give rise to smaller perturbations in the orbital elements. Some resonant orbits can be found, however, whose mean distance from the Earth is less than $4a_e$, but this can only occur with values of q greater than 2.

The angles $\eta_1, \eta_2, \ldots, \eta_{13}$ and $\zeta_1, \zeta_2, \ldots, \zeta_{12}$ are defined, in Table 6, in terms of quantities which have been given earlier. If any one of these angles is small, a resonance condition is obtained except for certain special values of q. The appropriate resonance condition in each case is given in the second column of Table 6. The quantities $\eta_1, \eta_2, \ldots, \eta_{13}$ apply when $p-q$ is even; while $\zeta_1, \zeta_2, \ldots, \zeta_{12}$ are applicable for odd values of $p-q$. The resonance conditions given in Section 4 correspond to η_3 and ζ_2.

When q takes the values 3, 4, 5, 6 or 7, one of the resonance conditions may be satisfied for orbits with a mean distance from the Earth's centre which is less than $4a_e$. For example, if $q = 4$ then η_7 will be small when

$$n = 4\omega_\gamma - 2\dot{\omega} - 4\dot{\Omega}.$$

A satellite with this mean motion will describe an orbit at a distance equal to about two and a half times the radius of the Earth. Thus these other resonance conditions will be useful when trying to determine high-order terms in the expression for the Earth's potential.

The equations for the derivatives of the potential with respect to the orbital elements are given in Tables 7 and 8. Since $\partial V_{pq}/\partial i$ is always divided by $\sin i$ in the formulae 1 (d),(e) for the time derivatives of Ω and ω, terms of order $\sin^3 i$, which occur when $p-q$ is even, have been included. Elsewhere terms of order greater than e, $e \sin i$ or $\sin^2 i$ have been neglected.

The first-order terms in Tables 7 and 8 correspond to orbits whose average radii are not less than $4a_e$. Any resonant orbits closer to the Earth, and these can only occur with values of q greater than 2, have secular terms of smaller order. It will be noticed, however, that $\partial V_{pq}/\partial e$ is divided by e in the formulae 1 (c), (e) for the time derivatives of χ and ω, and $\partial V_{pq}/\partial \chi$ is divided by e in the time derivative of e. Thus perturbations of order unity arise when η_5, ζ_6 and ζ_{10} are small. If values of q greater than 2 are being considered, resonant orbits derived from these quantities would have radii less than $4a_e$.

Table 6

$p-q$ even	Values of the mean motion for resonance
$\eta_1 = q(l+\beta_{pq})$	$\omega_r - \dot{\omega} - \dot{\Omega}$
$\eta_2 = q(l+\beta_{pq}) + M$	$[q\omega_r - q\dot{\omega} - q\dot{\Omega}]/(q+1)$
$\eta_3 = q(l+\beta_{pq}) - M$	$[q\omega_r - q\dot{\omega} - q\dot{\Omega}]/(q-1)$
$\eta_4 = q(l+\beta_{pq}) + 2M$	$[q\omega_r - q\dot{\omega} - q\dot{\Omega}]/(q+2)$
$\eta_5 = q(l+\beta_{pq}) - 2M$	$[q\omega_r - q\dot{\omega} - q\dot{\Omega}]/(q-2)$
$\eta_6 = q(l+\beta_{pq}) + M + 2(\omega+v)$	$[q\omega_r - (q+2)\dot{\omega} - q\dot{\Omega}]/(q+3)$
$\eta_7 = q(l+\beta_{pq}) - M - 2(\omega+v)$	$[q\omega_r - (q-2)\dot{\omega} - q\dot{\Omega}]/(q-3)$
$\eta_8 = q(l+\beta_{pq}) - M + 2(\omega+v)$	$[q\omega_r - (q+2)\dot{\omega} - q\dot{\Omega}]/(q+1)$
$\eta_9 = q(l+\beta_{pq}) + M - 2(\omega+v)$	$[q\omega_r - (q-2)\dot{\omega} - q\dot{\Omega}]/(q-1)$
$\eta_{10} = q(l+\beta_{pq}) + 2(\omega+v)$	$[q\omega_r - (q+2)\dot{\omega} - q\dot{\Omega}]/(q+2)$
$\eta_{11} = q(l+\beta_{pq}) - 2(\omega+v)$	$[q\omega_r - (q-2)\dot{\omega} - q\dot{\Omega}]/(q-2)$
$\eta_{12} = q(l+\beta_{pq}) + 4(\omega+v)$	$[q\omega_r - (q+4)\dot{\omega} - q\dot{\Omega}]/(q+4)$
$\eta_{13} = q(l+\beta_{pq}) - 4(\omega+v)$	$[q\omega_r - (q-4)\dot{\omega} - q\dot{\Omega}]/(q-4)$
$p-q$ odd	
$\zeta_1 = q(l+\beta_{pq}) + (\omega+v)$	$[q\omega_r - (q+1)\dot{\omega} - q\dot{\Omega}]/(q+1)$
$\zeta_2 = q(l+\beta_{pq}) - (\omega+v)$	$[q\omega_r - (q-1)\dot{\omega} - q\dot{\Omega}]/(q-1)$
$\zeta_3 = q(l+\beta_{pq}) + 3(\omega+v)$	$[q\omega_r - (q+3)\dot{\omega} - q\dot{\Omega}]/(q+3)$
$\zeta_4 = q(l+\beta_{pq}) - 3(\omega+v)$	$[q\omega_r - (q-3)\dot{\omega} - q\dot{\Omega}]/(q-3)$
$\zeta_5 = q(l+\beta_{pq}) + (\omega+v) + M$	$[q\omega_r - (q+1)\dot{\omega} - q\dot{\Omega}]/(q+2)$
$\zeta_6 = q(l+\beta_{pq}) - (\omega+v) - M$	$[q\omega_r - (q-1)\dot{\omega} - q\dot{\Omega}]/(q-2)$
$\zeta_7 = q(l+\beta_{pq}) - (\omega+v) + M$	$\omega_r - [(q-1)/q]\dot{\omega} - \dot{\Omega}$
$\zeta_8 = q(l+\beta_{pq}) + (\omega+v) - M$	$\omega_r - [(q+1)/q]\dot{\omega} - \dot{\Omega}$
$\zeta_9 = q(l+\beta_{pq}) + (\omega+v) + 2M$	$[q\omega_r - (q+1)\dot{\omega} - q\dot{\Omega}]/(q+3)$
$\zeta_{10} = q(l+\beta_{pq}) - (\omega+v) - 2M$	$[q\omega_r - (q-1)\dot{\omega} - q\dot{\Omega}]/(q-3)$
$\zeta_{11} = q(l+\beta_{pq}) - (\omega+v) + 2M$	$[q\omega_r - (q-1)\dot{\omega} - q\dot{\Omega}]/(q+1)$
$\zeta_{12} = q(l+\beta_{pq}) + (\omega+v) - 2M$	$[q\omega_r - (q+1)\dot{\omega} - q\dot{\Omega}]/(q-1)$

Table 7

p − q even

$$K = \frac{(-)^{\frac{1}{2}(p-q)}}{2^p}\frac{(p+q)!}{(\frac{1}{2}p-\frac{1}{2}q)!(\frac{1}{2}p+\frac{1}{2}q)!}\cdot\frac{J_{pq}}{a_e}\left(\frac{a_e}{a}\right)^{p+1}, \qquad x = p^2+p-q^2.$$

$$\frac{\partial V_{pq}}{\partial a} = -K\frac{(p+1)}{a}[\sin\eta_1 + A_1 e + A_2\sin^2 i],$$

$$A_1 = \tfrac{1}{2}(p+1)[\sin\eta_2+\sin\eta_3],$$

$$A_2 = \tfrac{1}{8}x[-2\sin\eta_1+\sin\eta_{10}+\sin\eta_{11}];$$

$$\frac{\partial V_{pq}}{\partial e} = K[B_1 + B_2 e + B_3\gamma^2 + B_4\sin^2 i],$$

$$B_1 = \tfrac{1}{4}[(p+2q+1)\sin\eta_2+(p-2q+1)\sin\eta_3],$$

$$B_2 = \tfrac{1}{4}[2p(p+1)\sin\eta_1+\{(p+1)(p+4)+2q(p+11)\}\sin\eta_4 + \\ +\{(p+1)(p+4)-2q(p+11)\}\sin\eta_5],$$

$$B_3 = \tfrac{1}{4}q[-\sin\eta_6+\sin\eta_7+\sin\eta_8-\sin\eta_9],$$

$$B_4 = \tfrac{1}{16}x[-2(p+2q+1)\sin\eta_2-2(p-2q+1)\sin\eta_3+(p+2q+5)\sin\eta_6+(p-2q+5)\sin\eta_7 + \\ +(p-2q-3)\sin\eta_8+(p+2q-3)\sin\eta_9];$$

$$\frac{\partial V_{pq}}{\partial \chi} = K[q\cos\eta_1 + C_1 e + C_2\gamma^2 + C_3\sin^2 i],$$

$$C_1 = \tfrac{1}{4}[(pq+3q+p+1)\cos\eta_2+(pq+3q-p-1)\cos\eta_3],$$

$$C_2 = -\tfrac{1}{4}q[\cos\eta_{10}+\cos\eta_{11}],$$

$$C_3 = \tfrac{1}{8}x[-2q\cos\eta_1+(q+2)\cos\eta_{10}+(q-2)\cos\eta_{11}];$$

$$\frac{\partial V_{pq}}{\partial \Omega} = K[q\cos\eta_1 + D_1 e + D_2\sin^2 i],$$

$$D_1 = \tfrac{1}{4}q(p+1)[\cos\eta_2+\cos\eta_3],$$

$$D_2 = \tfrac{1}{8}qx[-2\cos\eta_1+\cos\eta_{10}+\cos\eta_{11}];$$

$$\frac{\partial V_{pq}}{\partial \omega} = K[q\cos\eta_1 + E_1 e + E_2\gamma^2 + E_3\sin^2 i],$$

$$E_1 = \tfrac{1}{4}q(p+1)[\cos\eta_2+\cos\eta_3],$$

$$E_2 = -\tfrac{1}{4}q[\cos\eta_{10}+\cos\eta_{11}],$$

$$E_3 = \tfrac{1}{8}x[-2q\cos\eta_1+(q+2)\cos\eta_{10}+(q-2)\cos\eta_{11}];$$

$$\frac{\partial V_{pq}}{\partial i} = K\sin i[F_1\cos i + F_2\sec^3 i + F_3 e\cos i + F_4 e\sec^3 i + F_5\sin^2 i\sec^3 i],$$

$$F_1 = \tfrac{1}{4}x[-2\sin\eta_1+\sin\eta_{10}+\sin\eta_{11}],$$

$$F_2 = \tfrac{1}{4}q[-\sin\eta_{10}+\sin\eta_{11}],$$

$$F_3 = \tfrac{1}{8}x(p+1)[-2\sin\eta_2-2\sin\eta_3+\sin\eta_6+\sin\eta_7+\sin\eta_8+\sin\eta_9],$$

$$F_4 = \tfrac{1}{8}q(p+1)[-\sin\eta_6+\sin\eta_7-\sin\eta_8+\sin\eta_9],$$

$$F_5 = \tfrac{1}{32}qx[2\sin\eta_{10}-2\sin\eta_{11}-\sin\eta_{12}+\sin\eta_{13}].$$

Table 8

$p-q$ odd

$$K = \frac{(-)^{\frac{1}{2}(p-q-1)}}{2^p} \cdot \frac{(p+q+1)!}{(\frac{1}{2}p-\frac{1}{2}q-\frac{1}{2})!(\frac{1}{2}p+\frac{1}{2}q+\frac{1}{2})!} \cdot \frac{J_{pq}}{a_0}\left(\frac{a_0}{a}\right)^{p+1}, \qquad y = \tfrac{1}{4}(p^2+p-q^2-2).$$

$$\frac{\partial V_{pq}}{\partial a} = -K\frac{(p+1)}{a}\sin i[A_1+A_2 e],$$

$$A_1 = \tfrac{1}{4}[-\cos\zeta_1+\cos\zeta_2],$$

$$A_2 = \tfrac{1}{4}(p+1)[-\cos\zeta_3+\cos\zeta_6+\cos\zeta_7-\cos\zeta_8];$$

$$\frac{\partial V_{pq}}{\partial e} = K\sin i[B_1+B_2 e],$$

$$B_1 = \tfrac{1}{4}[-(p+2q+3)\cos\zeta_3+(p-2q+3)\cos\zeta_6+(p+2q-1)\cos\zeta_7-(p-2q-1)\cos\zeta_8],$$

$$B_2 = \tfrac{1}{8}[-2p(p+1)\cos\zeta_1+2p(p+1)\cos\zeta_2-(p^2+2pq+7p+22q+26)\cos\zeta_9+$$
$$+(p^2-2pq+7p-22q+26)\cos\zeta_{10}+(p^2+2pq+3p+22q-18)\cos\zeta_{11}-$$
$$-(p^2-2pq+3p-22q-18)\cos\zeta_{12}];$$

$$\frac{\partial V_{pq}}{\partial \chi} = K\sin i[C_1+C_2 e],$$

$$C_1 = \tfrac{1}{4}[(q+1)\sin\zeta_1-(q-1)\sin\zeta_2],$$

$$C_2 = \tfrac{1}{8}[(pq+2p+3q+4)\sin\zeta_3-(pq-2p+3q-4)\sin\zeta_6-(pq+3q-2)\sin\zeta_7+$$
$$+(pq+3q+2)\sin\zeta_8];$$

$$\frac{\partial V_{pq}}{\partial \Omega} = K\sin i[D_1+D_2 e],$$

$$D_1 = \tfrac{1}{4}q[\sin\zeta_1-\sin\zeta_2],$$

$$D_2 = \tfrac{1}{8}q(p+1)[\sin\zeta_3-\sin\zeta_6-\sin\zeta_7+\sin\zeta_8];$$

$$\frac{\partial V_{pq}}{\partial \omega} = K\sin i[E_1+E_2 e],$$

$$E_1 = \tfrac{1}{4}[(q+1)\sin\zeta_1-(q-1)\sin\zeta_2],$$

$$E_2 = \tfrac{1}{8}(p+1)[(q+1)\sin\zeta_3-(q-1)\sin\zeta_6-(q-1)\sin\zeta_7+(q+1)\sin\zeta_8];$$

$$\frac{\partial V_{pq}}{\partial i} = K[F_1\cos i+F_2 e\cos i+F_3\sin^2 i\cos i+F_4\sin^2 i\sec^2 i],$$

$$F_1 = \tfrac{1}{4}[-\cos\zeta_1+\cos\zeta_2],$$

$$F_2 = \tfrac{1}{8}(p+1)[-\cos\zeta_3+\cos\zeta_6+\cos\zeta_7-\cos\zeta_8],$$

$$F_3 = \tfrac{1}{8}y[3\cos\zeta_1-3\cos\zeta_2-\cos\zeta_3+\cos\zeta_4],$$

$$F_4 = \tfrac{1}{8}q[-\cos\zeta_1-\cos\zeta_2+\cos\zeta_3+\cos\zeta_4].$$

Numerical Tables Useful for Studying the Gravitational Field at Higher Elevations of the Earth

Chuji Tsuboi

"Now go, write it before them in a table and note it in a book."—*Isaiah* xxx, 8.

Summary

Expressions for the gravitational potential and force at a point (not necessarily on the Earth's axis) at any higher elevation of the Earth have been obtained in a simple summation series form in terms of zonal averages of surface gravity anomalies taken around the axis passing through the point. Numerical values to be multiplied to zonal averages of gravity anomalies in order to obtain the potential and force are given in tables. If the elevation is zero, our values for the potential naturally agree with those of Lambert's Q function.

If the distribution of gravity anomalies over the Earth's surface is given in a spherical harmonic series expression such as

$$\Delta g = \sum_{n=2}^{\infty} \sum_{s=0}^{\infty} \beta_{ns} S_{ns}, \tag{1}$$

the corresponding gravitational potential T and force F at the distance of r ($r > a$) from the Earth's centre are known to be given by

$$T = a \sum_{n=2}^{\infty} \sum_{s=0}^{\infty} \frac{\beta_{ns}}{n-1} \left(\frac{a}{r}\right)^{n+1} S_{ns}, \tag{2}$$

$$F = \sum_{n=2}^{\infty} \sum_{s=0}^{\infty} \beta_{ns} \left(\frac{a}{r}\right)^{n+2} S_{ns}. \tag{3}$$

Here, a is the radius of the Earth, S_{ns} a surface spherical harmonic of nth order and sth degree, and β_{ns} a constant coefficient. These are generalized statements of the theorem of Stokes which originally refers to the case $r = a$.

Various problems in this connection have been so fully discussed by many authors, including Sir Harold Jeffreys to whom this article is respectfully dedicated on the occasion of his seventieth birthday, that few of primary theoretical importance appear yet to be waiting for further investigations. Difficulties still lie,

however, in practical sides of the problem. No satisfactorily practical method has been presented by which to obtain required values of T and F at higher elevations numerically when gravity anomalies are known or given over the entire surface of the Earth.

Even in this age of electronic computors, formulae (2) and (3) are not necessarily the most convenient ones to be resorted to and some other formulae which are simpler to evaluate are desirable in many cases. The numerical table which was prepared by Walter D. Lambert & Frederic W. Darling (1936) is useful for the sea-level geoid. Enlargements of the table so as to make it applicable for higher elevations will be welcomed.

The object of the present article is twofold. One is to derive formulae for T and F which are easy to evaluate and the other to give a table of representative numerical values calculated therefrom together with an example of its use.

Let the circle in Figure 1 represent the Earth with the centre at C and O be the point at which the values of T and F are to be calculated. Draw a diameter of the

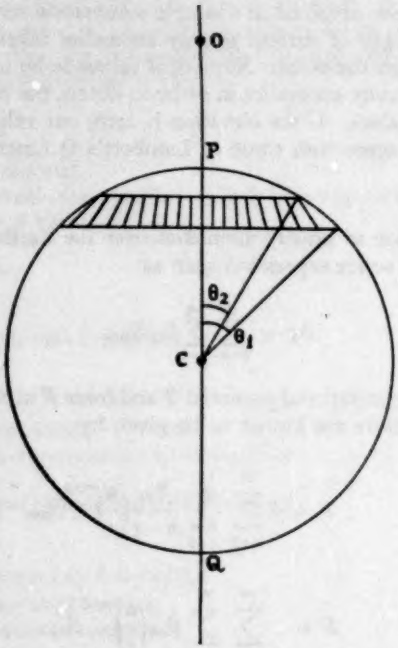


FIG. 1.

Earth which passes through O and let its intersections with the Earth's surface be P and Q. PQ can be any diameter and need not be the rotation axis of the Earth. T and F values at O are generally given in terms of gravity anomalies using a spherical harmonic series expression which may refer to any diametrical axis. But on the axis OC, the general expression in spherical harmonic series of S_{ns} is reduced to one in zonal harmonic series of P_n when that particular axis is referred

to. T and F become then

$$T = a \sum_{n=2}^{\infty} \frac{\beta_n}{n-1} \left(\frac{a}{r}\right)^{n+1} P_n, \tag{4}$$

$$F = \sum_{n=2}^{\infty} \beta_n \left(\frac{a}{r}\right)^{n+2} P_n. \tag{5}$$

This simplification is due to the circumstance that T and F values do not depend on how surface gravity anomalies are distributed with regard to azimuth within each of elementary spherical zones which extends from θ_1 to θ_2 as shown shaded in Figure 1. In fact, the values depend only on the azimuthal average of gravity anomalies taken in the zone.

The overall values of T and F are sums of contributions of such elementary zones. The contribution of an elementary zone which extends from θ_1 to θ_2 is the difference of the contributions of two spherical sectors, one extending from $\theta = \theta_1$ to 0, and the other from $\theta = \theta_2$ to 0. Thus our problem has been reduced to finding expressions for the gravitational potential T and force F due to a spherical sector which extends from $\theta = \theta$ to 0. By the reason stated already, gravity anomalies within this spherical sector may be taken to be uniformly $\overline{\Delta g}_m$.

Now any zonal distribution of a quantity $f(\theta)$ around the axis OC can be expressed in a zonal harmonic series, such as

$$f(\theta) = \sum_n \beta_n P_n$$

referring to that axis. β_n is a constant coefficient which is given by

$$\beta_n = \frac{2n+1}{2} \int_{-1}^{1} f(\theta) P_n \, d\mu,$$

where $\mu = \cos\theta$. In our special case that $f(\theta)$ is uniformly $\overline{\Delta g}_m$ from $\theta = \theta$ to 0 and is 0 from $\theta = \pi$ to θ, the integral taken from -1 to 1 may be replaced by that from μ to 1 and β_n becomes then to be

$$\beta_n = \frac{2n+1}{2} \overline{\Delta g}_m \int_{\mu}^{1} P_n \, d\mu. \tag{6}$$

But there is a relation

$$\int_{\mu}^{1} P_n \, d\mu = \frac{1}{2n+1} (P_{n-1} - P_{n+1}) \tag{7}$$

which simplifies matters greatly. In virtue of this relation, β_n becomes simply

$$\beta_n = \tfrac{1}{2} \overline{\Delta g}_m (P_{n-1} - P_{n+1}). \tag{8}$$

On inserting this relation into (4) and (5) and considering that all P_n's are unity on the axis, the following expressions are obtained:

$$T = a\overline{\Delta g}_m \sum_n \frac{1}{2(n-1)}\left(\frac{a}{r}\right)^{n+1}(P_{n-1}-P_{n+1}), \tag{9}$$

$$F = \overline{\Delta g}_m \sum_{n-2}^{\infty} \frac{1}{2}\left(\frac{a}{r}\right)^{n+2}(P_{n-1}-P_{n+1}), \tag{10}$$

which are easy to evaluate provided the values of P_n function are known.

In the case $r = a$, T is reduced to

$$T = a\overline{\Delta g}_m \sum_{n-2}^{\infty} \frac{1}{2(n-1)}(P_{n-1}-P_{n+1}) \tag{11}$$

and according to the terminology of Lambert and Darling our

$$\sum_{n-2}^{\infty} \frac{1}{2(n-1)}(P_{n-1}-P_{n+1})$$

is their Φ, which is

$$\Phi = \frac{1}{2}\left\{1+4\sin\frac{\theta}{2}-\cos\theta-6\sin^3\frac{\theta}{2}-\frac{7}{4}\sin^2\theta-\frac{3}{2}\sin^2\theta\log_e\left(\sin\frac{\theta}{2}+\sin^2\frac{\theta}{2}\right)\right\}$$

and these two must therefore be equal numerically. Although our new expression is much simpler in form than Φ, one might say that a significant difference between these two expressions lies in the fact that the former is not in a closed form while the latter is. But such a difference does not seem to the present writer to be very important practically, since the difference is scarcely more than a matter of formal nomenclature of a mathematical function.

As to the numerical table of P_n, there is an excellent one prepared by George C. Clark & Stuart W. Churchill (1957) which gives 6 place values of P_n for every degree from $n = 0$ up to $n = 80$. With the help of the table, the values of

$$\sum_{n-2}^{79} \frac{1}{2(n-1)}(P_{n-1}-P_{n+1})$$

were calculated for every 10 degrees. For the sake of reference, the summations were taken at the following 8 steps:

(1) $n = 2$ to $n = 10$
(2) $n = 2$ to $n = 20$
(3) $n = 2$ to $n = 30$
(4) $n = 2$ to $n = 40$
(5) $n = 2$ to $n = 50$
(6) $n = 2$ to $n = 60$
(7) $n = 2$ to $n = 70$
(8) $n = 2$ to $n = 79$

Table 1

$$\sum_{2}^{n} \frac{1}{2(n-1)}(P_{n-1}-P_{n+1})$$

θ / n	0°	10°	20°	30°	40°	50°
2 ~ 10	0·0000	0·2626	0·4351	0·5218	0·5351	0·5268
2 ~ 20	0·0000	0·2236	0·3895	0·5334	0·5573	0·5060
2 ~ 30	0·0000	0·2098	0·4060	0·5304	0·5629	0·5055
2 ~ 40	0·0000	0·1996	0·3956	0·5250	0·5640	0·5108
2 ~ 50	0·0000	0·2076	0·4001	0·5219	0·5633	0·5056
2 ~ 60	0·0000	0·2106	0·3981	0.5219	0·5620	0·5088
2 ~ 70	0·0000	0·2052	0·3981	0·5237	0·5591	0·5080
2 ~ 79	0·0000	0·2045	0·3997	0·5256	0·5607	0·5067
Lambert, Darling	0·0000	0·2068	0·3989	0·5241	0·5608	0·5078

60°	70°	80°	90°	100°	110°	120°
0·3795	0·1842	0·0205	−0·1492	−0·3124	−0·3787	−0·3889
0·3960	0·1981	0·0205	−0·1672	−0·2965	−0·3752	−0·3861
0·3935	0·2020	0·0136	−0·1601	−0·2900	−0·3749	−0·3762
0·3911	0·2038	0·0114	−0·1642	−0·2930	−0·3750	−0·3814
0·3936	0·2048	0·0133	−0·1615	−0·2996	−0·3752	−0·3820
0·3790	0·2054	0·0150	−0·1635	−0·2986	−0·3755	−0·3930
0·3809	0·2058	0·0144	−0·1619	−0·2974	−0·3758	−0·3920
0·3803	0·2062	0·0143	−0·1632	−0·2987	−0·3767	−0·3921
0·3806	0·2053	0·0138	−0·1626	−0·2981	−0·3764	−0·3928

130°	140°	150°	160°	170°	180°
−0·3598	−0·2812	−0·1690	−0·0869	−0·0283	−0·0000
−0·3493	−0·2782	−0·1754	−0·0879	−0·0233	−0·0000
−0·3556	−0·2763	−0·1782	−0·0862	−0·0220	−0·0000
−0·3528	−0·2750	−0·1789	−0·0878	−0·0231	−0·0000
−0·3532	−0·2742	−0·1783	−0·0864	−0·0235	−0·0000
−0·3541	−0·2739	−0·1776	−0·0875	−0·0234	−0·0000
−0·3528	−0·2739	−0·1773	−0·0866	−0·0228	−0·0000
−0·3535	−0·2749	−0·1779	−0·0867	−0·0229	−0·0000
−0·3534	−0·2745	−0·1778	−0·0870	−0·0230	−0·0000

Table 1 shows the way in which the sums converge to the corresponding values of Lambert and Darling, according as n is increased. From the satisfactory agreement of our values with theirs, it may be said that the summation up to $n = 79$ has fortunately been about sufficient and no further extension is much required.

Tables 2 and 3 give the values of

$$\Phi = \sum_{n=2}^{79} \frac{1}{2(n-1)}\left(\frac{a}{r}\right)^{n+1}(P_{n-1}-P_{n+1}),$$

and

$$\Psi = \sum_{n=2}^{79} \frac{1}{2}\left(\frac{a}{r}\right)^{n+2}(P_{n-1}-P_{n+1})$$

Table 2

$$\Phi = \frac{T}{a} = \sum_{n=2}^{79} \frac{1}{2(n-1)} \left(\frac{a}{r}\right)^{n+1} (P_{n-1} - P_{n+1}) \text{ and } \Delta\Phi$$

θ	$r = 1.0a$		$r = 1.1a$		$r = 1.2a$		$r = 1.3a$		$r = 1.4a$		$r = 1.5a$	
	Φ	$\Delta\Phi$	Φ	$\Delta\Phi$	Φ	$\Delta\Phi$	Φ	$\Delta\Phi$	Φ	$\Delta\Phi$	Φ	$\Delta\Phi$
0°	0·0000		0·0000		0·0000		0·0000		0·0000		0·0000	
		0·2068		0·1108		0·0665		0·0445		0·0310		0·0228
10	0·2068		0·1108		0·0665		0·0445		0·0310		0·0228	
		0·1921		0·1552		0·1180		0·0876		0·0670		0·0519
20	0·3989		0·2660		0·1845		0·1321		0·0980		0·0747	
		0·1252		0·1107		0·0924		0·0768		0·0628		0·0516
30	0·5241		0·3767		0·2769		0·2089		0·1608		0·1263	
		0·0367		0·0421		0·0421		0·0379		0·0342		0·0303
40	0·5608		0·4188		0·3190		0·2468		0·1950		0·1566	
		−0·0530		−0·0295		−0·0149		−0·062		−0·0016		0·0008
50	0·5078		0·3893		0·3041		0·2406		0·1934		0·1574	
		−0·1272		−0·0904		−0·0659		−0·0483		−0·0362		−0·0275
60	0·3806		0·2989		0·2382		0·1923		0·1572		0·1299	
		−0·1753		−0·1413		−0·1055		−0·0806		−0·0630		−0·0500
70	0·2053		0·1576		0·1327		0·1117		0·0942		0·0799	
		−0·1915		−0·1363		−0·1087		−0·0875		−0·0710		−0·0583
80	0·0138		0·0213		0·0240		0·0242		0·0232		0·0216	
		−0·1764		−0·1371		−0·1086		−0·0873		−0·0712		−0·0587
90	−0·1626		−0·1158		−0·0846		−0·0631		−0·0480		−0·0371	
		−0·1355		−0·1068		−0·0856		−0·0697		−0·0575		−0·0479
100	−0·2981		−0·2226		−0·1702		−0·1328		−0·1055		−0·0850	
		−0·0783		−0·0528		−0·0460		−0·0395		−0·0337		−0·0290
110	−0·3764		−0·2754		−0·2162		−0·1723		−0·1392		−0·1140	
		−0·0164		−0·0252		−0·0190		−0·0151		−0·0125		−0·0106
120	−0·3928		−0·3006		−0·2352		−0·1874		−0·1517		−0·1246	
		0·0394		0·0286		0·0209		0·0163		0·0126		0·0100
130	−0·3534		−0·2720		−0·2143		−0·1711		−0·1391		−0·1146	
		0·0789		0·0601		0·0472		0·0361		0·0300		0·0241
140	−0·2745		−0·2119		−0·1671		−0·1350		−0·1091		−0·0905	
		0·0967		0·0743		0·0586		0·0475		0·0376		0·0313
150	−0·1778		−0·1376		−0·1085		−0·0875		−0·0715		−0·0592	
		0·0908		0·0701		0·0546		0·0445		0·0363		0·0301
160	−0·0870		−0·0675		−0·0539		−0·0430		−0·0352		−0·0291	
		0·0640		0·0496		0·0397		0·0315		0·0259		0·0214
170	−0·0230		−0·0179		−0·0142		−0·0115		−0·0093		−0·0077	
		0·0230		0·0179		0·0142		0·0115		0·0093		0·0077
180	0·0000		0·0000		0·0000		0·0000		0·0000		0·0000	

Table 3

$$\Psi = \sum_{n=2}^{79} \frac{1}{2}\left(\frac{a}{r}\right)^{n+2}(P_{n-1} - P_{n+1}) \text{ and } \Delta\Psi$$

	$r = 1{\cdot}0a$		$r = 1{\cdot}1a$		$r = 1{\cdot}2a$		$r = 1{\cdot}3a$		$r = 1{\cdot}4a$		$= 1{\cdot}5a$	
θ	Ψ	$\Delta\Psi$	Ψ	$\Delta\Psi$	Ψ	$\Delta\Psi$	Ψ	$\Delta\Psi$	Ψ	$\Delta\Psi$	Ψ	$\Delta\Psi$
0°	0·0000		0·0000		0·0000		0·0000		0·0000		0·0000	
		0·9698		0·4277		0·1926		0·0984		0·0560		0·0346
10	0·9698		0·4277		0·1926		0·0984		0·0560		0·0346	
		−0·0877		0·1192		0·1441		0·1134		0·0819		0·0584
20	0·8821		0·5469		0·3367		0·2118		0·1379		0·0930	
		−0·1366		−0·0323		0·0185		0·0354		0·0371		0·0332
30	0·7455		0·5146		0·3552		0·2472		0·1750		0·1262	
		−0·1723		−0·0962		−0·0488		−0·0213		−0·0066		0·0009
40	0·5732		0·4184		0·3064		0·2259		0·1684		0·1271	
		−0·1919		−0·1263		−0·0825		−0·0537		−0·0350		−0·0229
50	0·3813		0·2921		0·2239		0·1722		0·1334		0·1042	
		−0·1938		−0·1358		−0·0963		−0·0688		−0·0497		−0·0364
60	0·1875		0·1563		0·1276		0·1034		0·0837		0·0678	
		−0·1787		−0·1298		−0·0955		−0·0711		−0·0537		−0·0409
70	0·0088		0·0265		0·0321		0·0323		0·0300		0·0269	
		−0·1494		−0·1110		−0·0835		−0·0627		−0·0494		−0·0388
80	−0·1406		−0·0845		−0·0514		−0·0304		−0·0194		−0·0119	
		−0·1094		−0·0839		−0·0638		−0·0508		−0·0391		−0·0311
90	−0·2500		−0·1684		−0·1152		−0·0812		−0·0585		−0·0430	
		−0·0642		−0·0492		−0·0398		−0·0318		−0·0257		−0·0208
100	−0·3142		−0·2176		−0·1550		−0·1130		−0·0842		−0·0638	
		−0·0191		−0·0173		−0·0150		−0·0129		−0·0110		−0·0095
110	−0·3333		−0·2349		−0·1700		−0·1259		−0·0952		−0·0733	
		−0·0208		−0·0125		−0·0074		−0·0043		−0·0025		−0·0014
120	−0·3125		−0·2224		−0·1626		−0·1216		−0·0927		−0·0719	
		0·0510		0·0351		0·0248		0·0179		0·0132		0·0099
130	−0·2615		−0·1873		−0·1378		−0·1037		−0·0795		−0·0620	
		0·0686		0·0484		0·0353		0·0263		0·0199		0·0153
140	−0·1929		−0·1389		−0·1025		−0·0774		−0·0596		−0·0467	
		0·0724		0·0519		0·0377		0·0286		0·0219		0·0171
150	−0·1205		−0·0870		−0·0648		−0·0488		−0·0377		−0·0296	
		0·0629		0·0458		0·0339		0·0253		0·0196		0·0153
160	−0·0576		−0·0412		−0·0309		−0·0235		−0·0181		−0·0143	
		0·0426		0·0303		0·0228		0·0173		0·0133		0·0105
170	−0·0150		−0·0109		−0·0081		−0·0062		−0·0048		−0·0038	
		0·0150		0·0109		0·0081		0·0062		0·0048		0·0038
180	0·0000		0·0000		0·0000		0·0000		0·0000		0·0000	

and their successive difference $\Delta\Phi$ and $\Delta\Psi$ at every 10 degrees for five values of r/a, namely 1·1, 1·2, 1·3, 1·4 and 1·5.

In Figures 2 and 3, Φ and Ψ values are graphically shown.

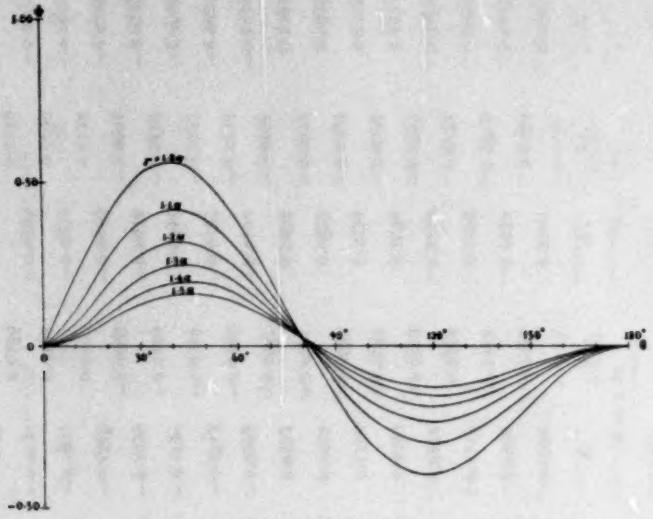


FIG. 2.—Φ Function.

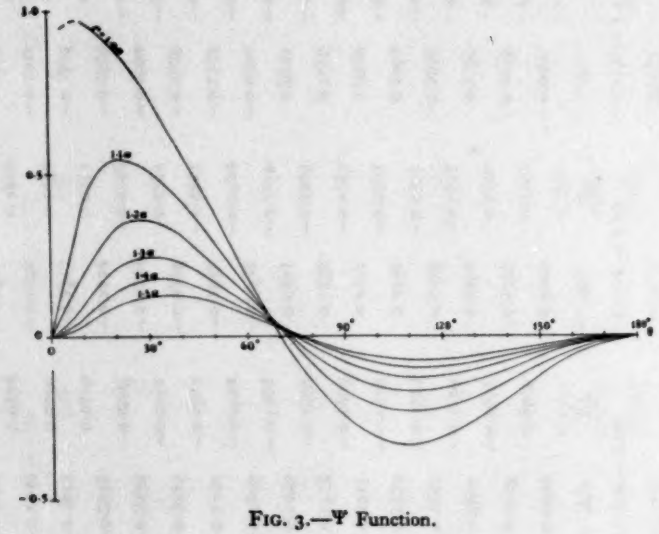


FIG. 3.—Ψ Function.

Now that the contributions of each elementary zone of unit gravity anomaly to T and F values at various higher elevations are numerically obtained, it will be interesting to calculate their actual values assuming a certain model distribution of

Table 4

$$\sum \overline{\Delta g}_m \Delta\Phi_m \text{ and } \sum \overline{\Delta g}_m \Delta\Psi_m$$

θ	Δg_m	$\overline{\Delta g}_m \cdot \Delta\Phi_m$						$\overline{\Delta g}_m \cdot \Delta\Psi_m$				
		$r = 1{\cdot}0a$	$r = 1{\cdot}1a$	$r = 1{\cdot}2a$	$r = 1{\cdot}3a$	$r = 1{\cdot}4a$	$r = 1{\cdot}5a$	$r = 1{\cdot}1a$	$r = 1{\cdot}2a$	$r = 1{\cdot}3a$	$r = 1{\cdot}4a$	$r = 1{\cdot}5a$
	mgal											
0°–10°	28	5·790	3·102	1·862	1·246	0·868	0·638	12·00	5·39	2·78	1·57	0·97
10–20	21	4·034	3·259	2·478	1·840	1·407	1·090	2·50	3·03	2·38	1·72	1·23
20–30	10	1·252	1·107	0·924	0·768	0·628	0·516	−0·32	0·18	0·35	0·37	0·33
30–40	−2	−0·073	−0·084	−0·084	−0·076	−0·068	−0·061	0·19	0·10	0·04	0·01	0·00
40–50	−9	0·477	0·266	0·134	0·056	0·014	−0·007	1·14	0·74	0·48	0·32	0·21
50–60	−11	1·399	0·994	0·725	0·531	0·398	0·303	1·49	1·06	0·76	0·55	0·40
60–70	−7	1·227	0·989	0·739	0·564	0·441	0·350	0·91	0·67	0·50	0·38	0·02
70–80	−1	0·192	0·136	0·109	0·088	0·071	0·058	0·11	0·08	0·06	0·05	0·69
80–90	5	−0·882	−0·686	−0·543	−0·437	−0·356	−0·294	−0·42	−0·32	−0·25	−0·20	−0·14
90–100	8	−1·084	−0·854	−0·685	−0·558	−0·460	−0·383	−0·40	−0·32	−0·25	−0·21	−0·17
100–110	6	−0·470	−0·317	−0·276	−0·237	−0·202	−0·174	−0·10	−0·09	−0·08	−0·07	−0·06
110–120	1	−0·016	−0·025	−0·019	−0·015	−0·013	−0·011	0·01	0·01	0·00	0·00	0·00
120–130	−4	−0·158	−0·114	−0·084	−0·065	−0·050	−0·040	−0·14	−0·10	−0·07	−0·05	−0·04
130–140	−6	−0·473	−0·361	−0·283	−0·217	−0·180	−0·145	−0·29	−0·21	−0·16	−0·12	−0·09
140–150	−5	−0·484	−0·372	−0·293	−0·238	−0·188	−0·157	−0·26	−0·19	−0·14	−0·11	−0·09
150–160	0	0·000	0·000	0·000	0·000	0·000	0·000	0·00	0·00	0·00	0·00	0·00
160–170	6	0·384	0·298	0·238	0·189	0·155	0·128	0·18	0·14	0·10	0·08	0·06
170–180	9	0·207	0·161	0·128	0·104	0·084	0·069	0·10	0·07	0·06	0·04	0·03
Sum		11·322	7·499	5·070	3·543	2·549	1·880	16·70	10·24	6·56	4·33	2·95

gravity anomalies over the Earth's surface. For this purpose, the following model adopted by Lambert and Darling (Example 2) will be taken.

θ	Δg_m
0–10°	28 mgal
10–20	21
20–30	10
30–40	−2
40–50	−9
50–60	−11
60–70	−7
70–80	−1
80–90	5
90–100	8
100–110	6
120–130	−4
130–140	−6
140–150	−5
150–160	0
160–170	6
170–180	9

The computed values of

$$\sum_m \overline{\Delta g}_m \overline{\Delta\Phi}_m \text{ and } \sum_m \overline{\Delta g}_m \Delta\Psi_m$$

for this model distribution at the five elevations are given in Table 4, together with Lambert's value which corresponds to the case $r = 1{\cdot}0a$.

It may have been noticed that in order to apply the present method to actual problems, zonal averages of surface gravity anomaly values taken around the axis passing through the point of interest must be known. It is admitted that this will mean an elaborate work to do unless some instrument is devised for making necessary averaging automatically.

Geophysical Institute,
Tokyo University,
Japan:
1960 *October.*

References

Lambert, W. D. & Darling, F. W., 1936. Tables for determining the form of the geoid and its indirect effect on gravity. *U.S. Coast and Geodetic Survey Sp. Publ.* No. 199.

Clark, G. C. & Churchill, S. W., 1957. *Tables of Legendre Polynomials.* (University of Michigan Press, Engineering Research Institute Publications.)

Intrinsic Coordinates in Practical Geodesy

Antonio Marussi

"We have advanced no vague reasoning about the well-known properties of spheroidical triangles and geodetical lines, in a case where, in fact, there is neither any such triangle nor any such line. . . ." *J. Ivory*, 1828.

Summary

The basic ideas that support intrinsic geodesy, i.e. the discipline which aims at the local description of the gravity field of the Earth by using only coordinates and quantities that have a physical reality and that are therefore accessible to actual observation, are recalled.

It is shown how the integrability conditions necessary for the existence of the coordinate surfaces, and the fundamental operators, e.g. the Christoffel symbols of the second kind connecting the principal trihedra of the intrinsic coordinate system, may be expressed in terms of the curvature parameters of the field, and of gravity.

An application of the theory is made to a classical geodetic problem, i.e. the generalized expansion of Legendre for the displacement of the potential (the dynamic height) along an optical path.

1. Man's natural environment forms part of a three-dimensional space. One of these three dimensions—loosely considered to be the vertical at a given place—has been instinctively differentiated from the other two; and hence it is not surprising that throughout geodesy—and particularly throughout practical geodesy—this third dimension has always been treated in a different manner.

As a consequence of this approach, which has profound roots in the human mind and in the historical development of geodesy, it has been necessary to use some reference surface, such as the geoid, which, although it possesses physical reality, is nevertheless inaccessible to direct observation; or like the ellipsoid and the spheroid, which are completely hypothetical. This approach thus also carries with it the burden of making complex and badly-defined reductions involving supplementary hypotheses—reductions which vitiate the accuracy with which our present-day observations can be made. Moreover, the arbitrary distinction of two dimensions from the third destroys that unity which is the true objective of scientific synthesis.

Thus it seemed desirable to develop a theory in which only those physical and geometrical quantities that are actually observed are used. Such a theory could restore to geodesy the unity which had been lost for the historical and psychological reasons explained.

At the Oslo Assembly of the I.U.G.G. in 1948 I presented the outline of a general theory which would permit the description of the Earth's gravity potential

field in terms of a system of coordinates intrinsic to the field itself. These coordinates are the astronomic latitude and longitude and the gravity potential (or dynamic height); and in the theory they are treated in a unified manner.

This theory, which I have subsequently elaborated in various papers, was originally intended to satisfy an aesthetic desire for unity, by returning the geodetic problem to its natural setting—three-dimensional space—and thereby avoiding the dichotomy between geometric and dynamic procedures, and unifying the disciplines of planimetry and altimetry. When it was first developed, the theory was not of immediate practical use, but the extremely rapid evolution both in scope and precision of geodetic measuring techniques now enables it to be applied to many problems of practical geodesy. The new techniques include the direct measurement of distance along geodesics in space, the use of targets at great altitudes such as rockets and satellites, and the facility with which gravity may now be measured both on land and at sea, and perhaps in the near future above the surface of the Earth.

There have already been successful applications of a unified theory. Brigadier Hotine, in Great Britain, has developed an elegant theory of spatial geodesy, in which the three dimensions are treated on an equal footing; and the theory involves all the quantities, and only those quantities, which are or could be the subject of geodetic observation. In the U.S.A., C. A. Whitten has applied this theory to practical problems. The ingenious method of stellar triangulation already developed by V. Väisälä in Finland also falls naturally into this scheme of ideas.

This is an appropriate place to observe that this geodetic revisionist movement —which has arisen in what we might call the *differential-geometric* approach to geodesy, and whose object is to gain knowledge of the actual external gravity field without debasing the measured quantities by the use of chimerical reductions— has a parallel which grew up almost simultaneously but quite independently in what we might call the *integral-dynamic* approach to geodesy—that dominated by Stokes's formula. Here, too, there has been an attempt to do away with the concept of the geoid—a surface whose theoretical definition is certainly unexceptionable, but which at least on the continents is in fact inaccessible—and to use instead the physical surface of the Earth as the surface needed for solving Stokes's problem, again avoiding those reduction terms which make the high precision of the measurements unusable.

This being the case, it seems appropriate to present here the fundamental equations of my theory in a form more suited to practical use.

2. We may consider two different systems of intrinsic coordinates in the external gravity field: first, a system of *general coordinates* comprising astronomic latitude ϕ, astronomic longitude λ, and gravity potential W; secondly, a system of *local Cartesian coordinates*, based on the three directions North, East, and Zenith, at any point. It is well known that the latter system cannot form the basis of any general system of coordinates. One system may be more convenient than the other according to the type of problem. When, for example, it is required to study scalar or vector quantities at geometrically interconnected points (as in triangulation or trilateration), the second system would be the more appropriate; whereas the first would be preferred when the quantities are to be considered at points that are not interconnected.

The fundamental differential operations are, for the first system, those of partial differentiation $\partial/\partial\phi$, $\partial/\partial\lambda$, $\partial/\partial W$ and for the second system, those of directional differentiation—$\partial/\partial s_1$ in the North direction, $\partial/\partial s_2$ in the East direction, and

$\partial/\partial h$ in the Zenith direction. These operators are related by the following equations, which are fundamental to the succeeding discussion:

$$\frac{\partial}{\partial s_1} = \kappa_1 \frac{\partial}{\partial \phi} + \frac{\tau}{\cos\phi}\frac{\partial}{\partial \lambda}$$

$$\frac{\partial}{\partial s_2} = \tau \frac{\partial}{\partial \phi} + \frac{\kappa_2}{\cos\phi}\frac{\partial}{\partial \lambda}$$

$$\frac{\partial}{\partial h} = \eta_1 \frac{\partial}{\partial \phi} + \eta_2 \frac{\partial}{\partial \lambda} - g\frac{\partial}{\partial W}.$$

Here κ_1 and κ_2 are the normal curvatures of the North and East lines, and τ is the geodetic torsion of the North line (equal and opposite to that of the East line). η_1 and η_2, the geometric components of the curvature-vector of the line of force along the North and East lines respectively, are given by

$$\eta_1 = \frac{\partial}{\partial s_1}\ln g; \quad \eta_2 = \frac{\partial}{\partial s_2}\ln g.$$

The conditions necessary for the existence of the coordinate surfaces $\phi = \text{const}$, $\lambda = \text{const}$, $W = \text{const}$, lead to five conditions for integrability. These may be most concisely written:

$$\tau\frac{\partial}{\partial \phi}\ln\frac{\tau\cos^2\phi}{K} + \frac{\kappa_2}{\cos\phi}\frac{\partial}{\partial \lambda}\ln\frac{\kappa_2}{K} = 0$$

$$\kappa_1\frac{\partial}{\partial \phi}\ln\frac{\kappa_1\cos\phi}{K} + \frac{\tau}{\cos\phi}\frac{\partial}{\partial \lambda}\ln\frac{\tau}{K} + \kappa_2\tan\phi = 0$$

$$\frac{\partial}{\partial W}\frac{\kappa_1}{K} = H_1\tan\phi - \frac{1}{\cos\phi}\frac{\partial H_2}{\partial \lambda}\frac{1}{g} = \left(\tan\phi\frac{\partial}{\partial \phi} - \frac{1}{\cos^2\phi}\frac{\partial^2}{\partial \lambda^2} - 1\right)\frac{1}{g}$$

$$\frac{\partial}{\partial W}\frac{\tau}{K} = \frac{\partial H_2}{\partial \phi} = \frac{1}{\cos\phi}\frac{\partial H_1}{\partial \lambda} + H_2\tan\phi = \frac{1}{\cos\phi}\left(\frac{\partial^2}{\partial \phi\partial \lambda} + \tan\phi\frac{\partial}{\partial \lambda}\right)\frac{1}{g}$$

$$\frac{\partial}{\partial W}\frac{\kappa_2}{K} = -\frac{\partial H_1}{\partial \phi}\frac{1}{g} = -\left(\frac{\partial^2}{\partial \phi^2} - 1\right)\frac{1}{g}.$$

To these may be added a sixth which is a form of Poisson's equation, and which may be written:

$$\frac{\partial}{\partial W}\ln g = \frac{1}{Kg}(\kappa_1\eta_2^2 - 2\tau\eta_1\eta_2 + \kappa_2\eta_1^2) + \frac{\kappa_1+\kappa_2}{g} + \frac{2\omega^2}{g^2}$$

$$= g(\kappa_1 H_1^2 + 2\tau H_1 H_2 + \kappa_2 H_2^2) + \frac{\kappa_1+\kappa_2}{g} + \frac{2\omega^2}{g^2}.$$

In the foregoing expressions $K = \kappa_1\kappa_2 - \tau^2$ is the Gaussian curvature of the equipotential surface $W = \text{const}$, g is the gravity and ω the angular velocity of rotation of the Earth. We may then put

$$H_1 = \frac{-1}{g}\frac{\partial}{\partial \phi}\ln g, \qquad H_2 = \frac{-1}{g\cos\phi}\frac{\partial}{\partial \lambda}\ln g,$$

where H_1 and H_2 are related to η_1 and η_2, as defined above, by the following expressions:

$$\eta_1 = -g(\kappa_2 H_1 + \tau H_2)$$
$$\eta_2 = -g(\tau H_1 + \kappa_1 H_2).$$

The first two conditions of integrability are concerned only with the parameters of curvature of the equipotential surface and their derivatives; whereas the remaining four express the rate of change of these parameters and of gravity on passing from one equipotential surface to another infinitely close.

The sets of three local Cartesian coordinates may be interconnected by means of Ricci's rotation coefficients, which are all expressed by the five parameters of curvature of the field κ_1, κ_2, τ, η_1 and η_2. Similarly, the bases vectors of the principal sets at the point P, $\partial P/\partial\phi$, $\partial P/\partial\lambda$ and $\partial P/\partial W$, relative to the general coordinates ϕ, λ, W, are related by the Christoffel symbols of the second kind. Allotting the numbers 1, 2, 3, to these coordinates, the Christoffel coefficients may be most clearly written

$$\left\{\begin{matrix}1\\11\end{matrix}\right\} = \frac{1}{K}\left(\kappa_2\frac{\partial}{\partial s_1}\ln\frac{\kappa_2}{Kg} - 2\tau^2\tan\phi\right)$$

$$\left\{\begin{matrix}1\\12\end{matrix}\right\} = \frac{-\cos\phi}{K}\left[\tau\frac{\partial}{\partial s_1}\ln\frac{\tau}{Kg} + \tau(\kappa_2-\kappa_1)\tan\phi\right]$$

$$\left\{\begin{matrix}1\\22\end{matrix}\right\} = \frac{\cos^2\phi}{K}\left[\kappa_1\frac{\partial}{\partial s_1}\ln\frac{\kappa_1}{Kg} + (\kappa_1\kappa_2+\tau^2)\tan\phi\right]$$

$$\left\{\begin{matrix}1\\13\end{matrix}\right\} = -\left(\frac{\partial}{\partial s_1}H_1 + \tau H_2\tan\phi + \frac{\kappa_1}{g}\right)$$

$$\left\{\begin{matrix}1\\23\end{matrix}\right\} = -\cos\phi\left(\frac{\partial}{\partial s_1}H_2 - \tau H_1\tan\phi + \frac{\tau}{g}\right)$$

$$\left\{\begin{matrix}1\\33\end{matrix}\right\} = \frac{1}{g}\left(\frac{\partial}{\partial s_1} + \eta_1\right)\frac{\partial}{\partial W}\ln g$$

$$\left\{\begin{matrix}2\\11\end{matrix}\right\} = \frac{1}{K\cos\phi}\left(\kappa_2\frac{\partial}{\partial s_2}\ln\frac{\kappa_2}{Kg} - 2\kappa_2\tau\tan\phi\right)$$

$$\left\{\begin{matrix}2\\12\end{matrix}\right\} = \frac{-1}{K}\left[\tau\frac{\partial}{\partial s_2}\ln\frac{\tau}{Kg} + (\kappa_2^2-\tau^2)\tan\phi\right]$$

$$\left\{\begin{matrix}2\\22\end{matrix}\right\} = \frac{\cos\phi}{K}\left(\kappa_1\frac{\partial}{\partial s_2}\ln\frac{\kappa_1}{Kg} + 2\kappa_2\tau\tan\phi\right)$$

$$\left\{\begin{matrix}2\\13\end{matrix}\right\} = \frac{-1}{\cos\phi}\left(\frac{\partial}{\partial s_2}H_1 + \kappa_2 H_2\tan\phi + \frac{\tau}{g}\right)$$

$$\left\{\begin{matrix}2\\23\end{matrix}\right\} = -\left(\frac{\partial}{\partial s_2}H_2 - \kappa_2 H_1\tan\phi + \frac{\kappa_2}{g}\right)$$

$$\left\{\begin{matrix}2\\33\end{matrix}\right\} = \frac{1}{g\cos\phi}\left(\frac{\partial}{\partial s_2} + \eta_2\right)\frac{\partial}{\partial W}\ln g$$

$$\begin{Bmatrix}3\\11\end{Bmatrix} = \frac{g\kappa_2}{K}; \quad \begin{Bmatrix}3\\12\end{Bmatrix} = \frac{-g\tau\cos\phi}{K}; \quad \begin{Bmatrix}3\\22\end{Bmatrix} = \frac{g\kappa_1\cos^2\phi}{K}$$

$$\begin{Bmatrix}3\\33\end{Bmatrix} = -\frac{\partial}{\partial W}\ln g; \quad \begin{Bmatrix}3\\13\end{Bmatrix} = \begin{Bmatrix}3\\23\end{Bmatrix} = 0.$$

If the field is symmetrical about the axis of rotation, all the preceding expressions become somewhat simplified. The derivatives with respect to the potential of the principal radii of curvature $1/\kappa_1$ and $1/\kappa_2$, and of the logarithm of gravity, may then be written:

$$\frac{\partial}{\partial W}\frac{1}{\kappa_1} = -\left(\frac{\partial^2}{\partial\phi^2}+1\right)\frac{1}{g}; \quad \frac{\partial}{\partial W}\frac{1}{\kappa_2} = \left(\tan\phi\frac{\partial}{\partial\phi}-1\right)\frac{1}{g}$$

$$\frac{\partial}{\partial W}\ln g = g\kappa_1\left(\frac{\partial}{\partial\phi}\frac{1}{g}\right)^2 + \frac{\kappa_1+\kappa_2}{g} + \frac{2\omega^2}{g^2}.$$

3. As a practical example, we shall apply the preceding to the generalized expansions of Legendre. The expansions allow the coordinates ϕ, λ, W (which we shall now denote by y^1, y^2, y^3 respectively, so that we can use the conventional notation of tensor calculus) to be displaced along a regular arc of any curve s in space, γ being the first curvature and θ the torsion. Hence

$$y^r = y_0^r + s\lambda_0^r + \frac{s^2}{2}\frac{d\lambda^r}{ds} + \dots$$

where λ^r are the controvariant components of the unit vector tangent to s. Hence by Frenet's formula

$$\frac{d\lambda^r}{ds} = \gamma\nu^r - \lambda^i\lambda^j\begin{Bmatrix}r\\ij\end{Bmatrix}$$

$$\frac{d^2\lambda^r}{ds^2} = \frac{d\gamma}{ds}\nu^r - \gamma(\gamma\lambda^r+\theta\beta^r) - 3\gamma\begin{Bmatrix}r\\ij\end{Bmatrix}\lambda^i\lambda^j + \left(2\begin{Bmatrix}r\\is\end{Bmatrix}\begin{Bmatrix}s\\jh\end{Bmatrix} - \frac{\partial}{\partial y^h}\begin{Bmatrix}r\\ij\end{Bmatrix}\right)\lambda^i\lambda^j\lambda^h$$

and so on, where ν^r and β^r are the controvariant components of the unit vector of the principal normal and of the binormal to s respectively.

The above formulae allow these expansions to be given explicitly; in particular λ^r has the form

$$\begin{aligned}\lambda^1 &= \kappa_1\sin z\cos\alpha + \tau\sin z\sin\alpha + \eta_1\cos z\\ \cos\phi\lambda^2 &= \tau\sin z\cos\alpha + \kappa_2\sin z\sin\alpha + \eta_2\cos z\\ \lambda^3 &= -g\cos z\end{aligned}$$

where $\bar\alpha$ and $\bar z$ are the azimuth and the zenith distance of the tangent at s.

By analogy

$$\nu^1 = \kappa_1\sin\bar z\cos\bar\alpha + \tau\sin\bar z\sin\bar\alpha + \eta_1\cos\bar z.$$

In this case $\bar\alpha$ and $\bar z$ are the azimuth and zenith distance of the principal normal to s, related to α and z by the orthogonality condition

$$\tan z\tan\bar z\cos(\alpha-\bar\alpha)+1 = 0.$$

If the curve s of displacement is an optical path, we can assume without appreciable error that its osculating plane is everywhere vertical, in which case $\alpha = \bar\alpha$,

and $\bar{z} = z + 90°$. Here we shall confine ourselves to giving the explicit expression for the displacement of potential (or dynamic height) along the curve s under consideration (which we may regard as an optical path). By a somewhat lengthy but simple calculation

$$W = W_0 - sg_0 \cos z_0 + \frac{s^2 g_0}{2}\left[\gamma \sin z_0 - \kappa_\alpha \sin^2 z_0 - \sin 2z_0 \frac{\partial}{\partial s_\alpha} \ln g + \right.$$
$$\left. + \left(\kappa_1 + \kappa_2 + \frac{2\omega^2}{g}\right)\cos^2 z\right] + \frac{s^3 g}{6}\left(\frac{d\gamma}{ds} + \ldots\right) + \ldots .$$

Here z is the observed zenith distance at the near end (without correction for refraction); κ_α is the normal curvature of the equipotential surface in the vertical plane containing s, and obviously

$$\frac{\partial}{ds_\alpha} = \cos\alpha \frac{\partial}{ds_1} + \sin\alpha \frac{\partial}{ds_2}.$$

It is obvious that the value of each coefficient should be calculated for the near end.

The preceding formula does not contain products of the curvatures and their derivatives, but retains the terms in $d\gamma/ds$. For displacement of zenith distance we have the expression

$$\cos z = \frac{g_0}{g}\cos z_0 - s\left[\gamma \sin z_0 - \kappa_\alpha \sin^2 z_0 - \sin 2z_0 \frac{\partial}{\partial s_\alpha} \ln g + \right.$$
$$\left. + \left(\kappa_1 + \kappa_2 + \frac{2\omega^2}{g}\right)\cos^2 z_0\right] + \frac{s^2}{2}\left(\frac{d\gamma}{ds}\sin z_0 + \ldots\right) + \ldots .$$

This formula requires the values of gravity at both ends. For the displacement of the gravity value in the system of general coordinates, we may use the ordinary Taylor's expansion

$$\ln g = \left(1 + \Delta y^i \frac{\partial}{\partial y^i} + \ldots\right) \ln g_0.$$

In the local Cartesian system the expansion

$$\ln g = \left(1 + s\frac{d}{ds} + \ldots\right) \ln g_0$$

may be used. In both these expansions the equations given above allow all the derivatives concerned to be expressed in terms of the surface derivatives alone.

This example shows how the theory may be very easily applied to give the most rigorous solution to the problem of evaluating potential differences—a problem occurring in the most elementary surveying. In a like manner these procedures may be applied to more complex geodetic calculations, such as the displacement of geographic coordinates along any curve in space, with little more difficulty than by classical methods, but with the advantage of providing results notably more concordant with physical reality.

The University,
Trieste:
1960 *October.*

References

Hotine, M., 1957. Metrical properties of the Earth's gravitational field and geodetic coordinate systems. A.I.G., Toronto, 1957.

Hotine, M., 1957. A primer of non-classical geodesy. A.I.G., Toronto, 1957.

Hotine, M., 1960. The third dimension in geodesy. A.I.G., Helsinki, 1960.

Ivory, J., 1828. On the method in the trigonometrical survey for finding the difference of longitude of two stations very little different in latitude. *The Philosophical Magazine and Annals of Philosophy* (London, Dec. 1828), 432.

Marussi, A., 1950. Sviluppi di Legendre generalizzati per una curva qualunque dello spazio. *Rend. Acc. Naz. Lincei*, Serie VIII, Vol. IX, Fasc. 1–2. (Roma, 1950.)

Marussi, A., 1951. Fondamenti di Geodesia Intrinseca. *Pubblicazioni della Comm. Geod. Italiana*, Serie III, Memoria no. 7. (Milano, 1951.)

Väisälä, Y., 1947. An astronomical method of triangulation. *Finn. Acad. Sc.*, 1946, 99–107. (Helsinki, 1947.)

II. Seismology — Body Waves

	PAGE
K. E. BULLEN, Seismic ray theory	93
M. BÅTH, Polarization of transverse seismic waves	106
I. LEHMANN, *S* and the structure of the upper mantle	124
P. CALOI, Seismic waves from the outer and the inner core	139
F. GASSMANN, Solution of an *n*-layer problem by a seismic reflection method	151

Seismic Ray Theory

K. E. Bullen

"Touch'd with an interior ray."—*Richard Crashaw, To the Name above every Name.*

Summary

Seismic ray theory is developed *ab initio* (apart from one or two standard elementary results) with special emphasis on the variables ζ and α, where $\zeta = d \log v/d \log r$ and $\alpha = 2/(1-\zeta)$. Travel-time–distance relations are examined for a variety of types of velocity distribution, including the case α constant and cases where α and v change discontinuously, or rapidly but continuously, with increase of depth. The analysis is designed to provide an improved basis for working out in quantitative detail the effect on travel-times of the various types of velocity variation likely to be relevant to the Earth. In particular, it is hoped that the analysis will lead to the most effective use of ray theory in the current difficult problems of the structure of the Earth's outer mantle.

Advantage has been taken of the opportunity to present a number of previous results involving ζ and α in revised form, as part of a wider logical development. Previous work on deriving seismic velocity distributions from travel-time data is generalized.

The aim of the paper has been to set down a terse account of the basic theory, and no numerical applications have been included. References are given, however, to papers containing applications.

1. Preliminaries

In earlier work of the writer (1945, 1955) on seismic ray theory, the variables ζ and α, defined below, arose incidentally. More specific use of ζ and α has been made in recent papers (1960, 1960a). The purpose of the present paper is to restate ray theory for a spherically symmetrical Earth in a form in which central use is made of these variables. Previous results will be collected together and presented with revisions and extensions. It is hoped that the presentation may serve as a succinct introduction to the use of ray theory in current problems on the structure of the Earth's interior.

1.1 Immediate properties of ζ and α

The variables η, ζ and α are defined as functions of the distance r from the centre of a (spherically symmetrical) Earth model by

$$\eta = r/v, \tag{1}$$

$$\zeta = \frac{d\log v}{d\log r}, \tag{2}$$

$$\alpha = \frac{2}{1-\zeta} = \frac{2d\log r}{d\log \eta}, \tag{3}$$

where v denotes the appropriate seismic wave velocity.

ζ and α derive immediate significance from the fact that the simple condition

$$\zeta < 1, \quad \text{or} \quad \alpha > 0, \tag{4}$$

(or $d\eta/dr > 0$), determines whether or not there can exist a ray with lowest point at a given level r. (This follows from (2), since $v^{-1}\,dv/dr$ is the downward curvature of a ray at r (Bullen, 1954, p. 87), and r^{-1} is the curvature of the level-surface.) Values of r at which ζ and α pass through the values 1 and ∞, respectively, are associated with discontinuous changes in shapes of rays of a family; correspondingly, divergent integrals appear in the mathematical ray analysis. The development to follow incidentally illustrates the power of analytical processes which rest on variables well based physically or geometrically, as are ζ and α.

In parts of the Earth where v decreases with the depth, but not so rapidly as to violate the condition $dv/dr < v/r$, we have $0 < \zeta < 1$ and $2 < \alpha < \infty$. Thus ζ and α also lend themselves to testing suspected regions of negative velocity gradient.

1.2 Basic relations between t, Δ and p

Except where otherwise indicated, each seismic ray of a family will have both end-points on a given level-surface $r = r_0$. The angular length Δ of a ray, the travel-time t of a pulse along it, and the usual ray-parameter p are (Bullen, 1959, p. 109; an interesting alternative derivation is given in Bullen, 1954, p. 83) connected by

$$p = dt/d\Delta. \tag{5}$$

Subscripts 0 and Δ will denote values of functions of r at $r = r_0$ and at the lowest point of the ray of length Δ, respectively.

Let e denote the angle between a ray and a level-surface of radius r. Then Snell's law (1959, p. 108) gives

$$p = \eta \cos e = \eta_\Delta; \qquad p_{\Delta=0} = \eta_0. \tag{6, 7}$$

The relations $\cot e = r\,d\theta/dr$, $\sin e = dr/ds$ (where s denotes arc-length and θ angular coordinate along a ray), and $t = \int ds/v$, lead by (6) to the basic relations

$$\Delta = 2p \int_{r_\Delta}^{r_0} r^{-1}(\eta^2 - p^2)^{-\frac{1}{2}}\,dr, \tag{8}$$

$$t = 2 \int_{r_\Delta}^{r_0} \eta^2 r^{-1}(\eta^2 - p^2)^{-\frac{1}{2}}\,dr. \tag{9}$$

The further relation

$$t - p\Delta = 2 \int_{r_\Delta}^{r_0} r^{-1}(\eta^2 - p^2)^{\frac{1}{2}}\,dr, \tag{10}$$

readily reducible from (5), (8) and (9), has been much used in numerical applications of ray theory, since the integrand on the right side remains finite as $\eta \to \eta_\Delta$.

1.3 Introduction of dimensionless variables

It is convenient to introduce dimensionless variables x, μ, λ and T, where

$$x = r/r_0; \quad \mu = \eta/\eta_0; \quad \lambda = p/\eta_0; \quad T = t/\eta_0. \quad (11, 12, 13, 14)$$

Thus

$$x_0 = \mu_0 = 1; \quad \lambda_{\Delta=0} = 1; \quad \lambda = dT/d\Delta = \mu_\Delta. \quad (15, 16, 17)$$

The variables x, μ, ζ and α are all functions of r. But η_Δ, $\mu\Delta$, ζ_Δ and α_Δ, like T, Δ and λ (as well as θ and ψ, to be defined in Section 2, and e_0) are functions of Δ (or λ).

When variation from ray to ray of a family is being considered, the argument λ will be taken to decrease continuously, starting from the value unity for which Δ and e_0 are both zero.

In dimensionless variables, (8), (9) and (10) become

$$\Delta = 2\lambda \int_{x_\Delta}^{1} x^{-1}(\mu^2-\lambda^2)^{-\frac{1}{2}}\,dx \quad (18)$$

$$= \int_{\lambda}^{1} \alpha\lambda\mu^{-1}(\mu^2-\lambda^2)^{-\frac{1}{2}}\,d\mu, \quad (19)$$

$$T = \int_{\lambda}^{1} \alpha\mu(\mu^2-\lambda^2)^{-\frac{1}{2}}\,d\mu, \quad (20)$$

$$T-\lambda\Delta = \int_{\lambda}^{1} \alpha\mu^{-1}(\mu^2-\lambda^2)^{\frac{1}{2}}\,d\mu. \quad (21)$$

2. Variation of T, Δ, with λ; the functions of θ and ψ

Integrating (19) and (20) by parts gives

$$\Delta = [\alpha\cos^{-1}(\lambda/\mu)]_\lambda^1 - \int_{\alpha_\Delta}^{\alpha_0} \cos^{-1}(\lambda/\mu)\,d\alpha$$

$$= \alpha_0\cos^{-1}\lambda - \int_{\alpha_\Delta}^{\alpha_0} \cos^{-1}(\lambda/\mu)\,d\alpha; \quad (22)$$

$$T = \alpha_0(1-\lambda^2)^{\frac{1}{2}} - \int_{\alpha_\Delta}^{\alpha_0} (\mu^2-\lambda^2)^{\frac{1}{2}}\,d\alpha. \quad (23)$$

(Treating (21) similarly yields only a trivial combination of (22) and (23). Thus forms corresponding to (10) and (21) are not needed in the present treatment.)

Differentiating (22) and (23) and using (17) gives

$$\frac{d\Delta}{d\lambda} = -\alpha_0(1-\lambda^2)^{-\frac{1}{2}} + \int_{\alpha_\Delta}^{\alpha_0} (\mu^2-\lambda^2)^{-\frac{1}{2}}\, d\alpha, \tag{24}$$

$$\frac{dT}{d\lambda} = -\alpha_0\lambda(1-\lambda^2)^{-\frac{1}{2}} + \int_{\alpha_\Delta}^{\alpha_0} \lambda(\mu^2-\lambda^2)^{-\frac{1}{2}}\, d\alpha. \tag{25}$$

((24) is equivalent to the result expressed in equations (17), (18) and (19) of the writer's text (1959, page 112), but the present choice of variables makes the derivation very much simpler. (25) comes, alternatively, directly from (17) and (24).)

Let θ and ψ denote the integrals in (24) and (25), so that

$$\theta = \int_{\alpha_\Delta}^{\alpha_0} (\mu^2-\lambda^2)^{-\frac{1}{2}}\, d\alpha, \tag{26}$$

$$\psi = \int_{\alpha_\Delta}^{\alpha_0} \lambda(\mu^2-\lambda^2)^{-\frac{1}{2}}\, d\alpha. \tag{27}$$

Thus

$$\theta = \frac{d\Delta}{d\lambda} + \alpha_0(1-\lambda^2)^{-\frac{1}{2}} = \frac{d}{d\lambda}(\Delta - \alpha_0 \cos^{-1}\lambda); \tag{28}$$

$$\psi = \frac{dT}{d\lambda} + \alpha_0\lambda(1-\lambda^2)^{-\frac{1}{2}} = \frac{d}{d\lambda}(T - \alpha_0(1-\lambda^2)^{\frac{1}{2}}) = \lambda\theta. \tag{29}$$

By (17), (28) and (29), θ and ψ are known functions of λ when T is a known function of Δ.

Also, just as (8) and (9) can be used to determine t and Δ as functions of p (or λ) when a velocity distribution is given, so can (26) and (27) be used to determine θ and ψ in terms of λ.

3. Derivation of velocity distribution from travel-time data

When t is given as a function of Δ, the problem of determining v as a function of r involves reversing the process mentioned in the last paragraph of Section 2. In the classical procedure, r is determined as a function of η by solving (8) as an integral equation. (This corresponds to solving (18) to find x as a function of μ.)

By a slight generalization of the classical theory, it is possible to include alternative procedures which involve the variable α.

3.1 Lemma

The generalization rests on the lemma:
Given that

$$f(\lambda) = \int_{\lambda}^{1} (\mu^2-\lambda^2)^{-\frac{1}{2}}\phi'(\mu)\,d\mu, \tag{30}$$

where f and ϕ and their derivatives are continuous in the ranges involved, then

$$\tfrac{1}{2}\pi(\phi_0-\phi_1) = \int_{\mu_1}^{1} \lambda f(\lambda)(\lambda^2-\mu_1^2)^{-\frac{1}{2}}\,d\lambda, \tag{31}$$

where the subscripts 0 and 1 correspond to $\mu = 1$ and to an arbitrarily chosen value of μ inside the allowed range, respectively.

The lemma is a particular case of the theorem which expresses the solution of Abel's integral equation, but is more general than that required for example in Rasch's process of deriving velocities (Jeffreys 1939). With a slight change of variables, the lemma is (according to Bôcher, 1914, p. 10) actually equivalent to Abel's main result in this field, Abel having not considered the more general integral equation which in fact bears his name.

The lemma may be simply proved by applying the operation

$$\int_{\mu_1}^{1} \lambda(\lambda^2-\mu_1^2)^{-\frac{1}{2}}\,d\lambda$$

to both sides of (20), changing the order of integration, and using the well-known result that

$$\int_{\mu_1}^{\mu} \{(\lambda^2-\mu_1^2)(\mu^2-\lambda^2)\}^{-\frac{1}{2}}\lambda\,d\lambda = \tfrac{1}{2}\pi.$$

In our applications, λ and μ will be given the same meanings as in Section 1.3.

3.2 Application of lemma to solving (18); (derivation of classical formula for determining v as a function of r)

In (30), put, in particular, $f = \Delta/2\lambda$ and $\phi = \ln x$; this gives (18). By (31), the solution is

$$\begin{aligned}\pi \ln x_1 &= \int_{1}^{\mu_1} \Delta(\lambda^2-\mu_1^2)^{-\frac{1}{2}}\,d\lambda \\ &= [\Delta \cosh^{-1}(\lambda/\mu_1)]_1^{\mu_1} - \int_{0}^{\Delta_1} \cosh^{-1}(\lambda/\mu_1)\,d\Delta \\ &= \int_{\Delta_1}^{0} \cosh^{-1}(\lambda/\mu_1)\,d\Delta, \end{aligned} \tag{32}$$

by (16), where Δ_1 here denotes the length of the ray for which $\lambda = \mu_1$.

(32) is equivalent to the classical formula (Bullen, 1959, p. 119, equation (30)) for deriving v as a function of r when t is a given function of Δ. It connects an arbitrarily assigned μ_1 with the corresponding x_1, and thus, by (1), (11) and (12), v_1 with r_1.

3.3 Application to solving (26)

In (30), put $f = \theta$ and $\phi = \alpha$, giving (26). By (31), the solution is

$$\begin{aligned}\tfrac{1}{2}\pi(\alpha_1-\alpha_0) &= \int_1^{\mu_1} \theta\lambda(\lambda^2-\mu_1^2)^{-\frac{1}{2}}\,d\lambda \\ &= [\theta(\lambda^2-\mu_1^2)^{\frac{1}{2}}]_1^{\mu_1} - \int_0^{\theta_1} (\lambda^2-\mu_1^2)^{\frac{1}{2}}\,d\theta \\ &= \int_{\theta_1}^{0} (\lambda^2-\mu_1^2)^{\frac{1}{2}}\,d\theta, \qquad (33)\end{aligned}$$

since $\alpha_\Delta = \alpha_0$ when $\lambda = 1$, and (26) then gives $\theta = 0$.

(33) enables α_1 to be connected with an assigned μ_1 when θ is a known function of λ, and hence again, by (1) and (3), v_1 with r_1.

3.4 Application to solving (27)

In (30), put $f = \psi/\lambda$ and $\phi = \alpha$, giving (27). This case is similar to that of Section 3.2, except that ψ replaces $\frac{1}{2}\Delta$ and α replaces $\ln x$. Further, by (29), ψ vanishes when θ vanishes, and therefore when $\lambda = 1$. Hence the solution is

$$\tfrac{1}{2}\pi(\alpha_1-\alpha_0) = \int_{\psi_1}^{0} \cosh^{-1}(\lambda/\mu_1)\,d\psi. \qquad (34)$$

(34) again connects α_1 with μ_1, and may be used as an alterative to (33).

4. The particular velocity distribution given by ζ = constant

If ζ is constant, $\alpha = \alpha_0$ for all r or x. By (26) and (27), θ and ψ are then zero so that

$$\frac{d\Delta}{d\lambda} = -\alpha_0(1-\lambda^2)^{-\frac{1}{2}}, \qquad (35)$$

$$\frac{dT}{d\lambda} = -\alpha_0\lambda(1-\lambda^2)^{-\frac{1}{2}}; \qquad (36)$$

and hence

$$\Delta = \alpha_0\cos^{-1}\lambda, \qquad (37)$$

$$T = \alpha_0\sqrt{(1-\lambda^2)} \qquad (38)$$

$$= \alpha_0\sin(\Delta/\alpha_0). \qquad (39)$$

This particular case is highly important because of: (i) the simplicity of the (T, Δ, λ) relations (37)–(39); (ii) the simplicity of the corresponding velocity law

$$\alpha = \alpha_0, \text{ or } \zeta = \zeta_0, \text{ or } v = v_0 x^{\zeta_0}; \tag{40}$$

(iii) the fact that (39) gives a close representation of actual (T, Δ) data in various parts of the Earth.

It will be noticed that the case ζ = constant includes the much more special case v = constant (for which $\zeta_0 = 0$, $\alpha_0 = 2$).

In many applications, ζ_0 is negative, corresponding (by (2)) to $(dv/dr)_0$ being negative. Then, (40) formally gives $v = \infty$ at $x = 0$. But if (40) is taken as a first approximation to an actual velocity distribution, the effects of this singularity can be avoided by using the form $\alpha = \alpha_0$ instead of the form in v.

The variables θ and ψ derive importance from the fact that, by (28) and (29), they vanish when (37)–(40) hold. And when (37)–(40) do not hold, either of (33) and (34) enables the necessary corrections to (40) to be deduced.

The use of θ and ψ, with equations (33) and (34), in deriving velocity distributions is illustrated elsewhere (1960) in some detail.

4.1 Use of the angle of emergence e_0

By (6) and (13), $\lambda = \cos e_0$, enabling (37) and (38) to be written very simply as

$$\Delta = \alpha_0 e_0, \qquad T = \alpha_0 \sin e_0; \tag{41}$$

(35) and (36) are given by (41) and

$$de_0/d\lambda = -\operatorname{cosec} e_0. \tag{42}$$

Because of the simplicity of (41), there are advantages in using e_0 in place of λ in certain contexts. (See e.g. Lehmann, 1959, p. 403.) The advantages are, however, limited by the fact that the usual seismic data give numerical values of t, Δ and λ fairly directly in practice, and λ (or p) is for many purposes the most convenient argument to use. This applies especially in many cases of deviation from (39), including those discussed in Sections 5 and 6.

4.2 Segments of rays; focal-depth allowance

The formulae (35)–(39) apply to rays terminating at the level $r = r_0$. We now consider a segment of a ray, taken as lying wholly on one side of the ray's lowest point, and terminating at the levels r_1 and r_2, say.

Let Δ_{12} and T_{12} apply to such a segment, and Δ_1 and Δ_2 to segments extending from the levels r_1 and r_2, respectively, to the lowest point. (Δ_1 has here a different meaning from that in Section 3.2.) Then, from (37), noting that $\lambda = \lambda/\mu_0$, we derive

$$\Delta_{12} = \Delta_1 - \Delta_2 = \tfrac{1}{2}\alpha_0\{\cos^{-1}(\lambda/\mu_1) - \cos^{-1}(\lambda/\mu_2)\}. \tag{43}$$

Similarly,

$$T_{12} = \tfrac{1}{2}\alpha_0\{\sqrt{(\mu_1^2 - \lambda^2)} - \sqrt{(\mu_2^2 - \lambda^2)}\}. \tag{44}$$

(A relation corresponding to (43) and (44) has been obtained in a quite different manner by Jeffreys (1959).)

If, in the right sides of (43) and (44), μ_1 is replaced by unity, the formulae give the focal-depth allowances for Δ and T for a ray starting from a focus at the level r_2.

The contributions to Δ, T and $d\Delta/d\lambda$ for the pair of segments of a whole ray which lie between the levels r_0 and r_1 are given respectively by

$$\alpha_0\{\cos^{-1}\lambda - \cos^{-1}(\lambda/\mu_1)\}, \tag{45}$$

$$\alpha_0\{(1-\lambda^2)^{\frac{1}{2}} - (\mu_1^2-\lambda^2)^{\frac{1}{2}}\}, \tag{46}$$

$$\alpha_0\{-(1-\lambda^2)^{-\frac{1}{2}} + (\mu_1^2-\lambda^2)^{-\frac{1}{2}}\}. \tag{47}$$

The formulae (45)–(47) are relevant when α has a constant value down to the level r_1 and then changes. The formulae are important auxiliaries to the discussion in the next section on discontinuities.

5. Effects of sudden or rapid changes in the velocity v

We shall now consider models in which the representation given by (37)–(40) is satisfactory in an uppermost region L for which $r_0 \leqslant r \leqslant r_1$, and in which strong changes in v as a function of r set in at the level r_1. Such models are of practical importance because of the various types of behaviour of v that are met with in parts of the Earth.

We let α have the constant (positive) value α_0 throughout L.

All rays to be considered in Section 5 will have both end-points at the level r_0. The formulae of the last section enable the results to be adapted to cases of focal depth.

Subscripts 0, 1 and 2 will denote values at the levels r_0, r_1 and r_2. For functions which are discontinuous at r_1, the subscripts a and b will denote values just above and just below this level.

5.1 Effect of discontinuous increase in velocity gradient; velocity continuous

Let the layer N, in which we take α to have the constant value $k\alpha_0$, occupy the region immediately below L for some distance. We here take the velocity gradient $(-dv/dr)$ to jump discontinuously at the level r_1 between L and N, v itself being continuous. Thus $\alpha_b < \alpha_a$, and $k < 1$, while $\mu_a = \mu_b = \mu_1$.

For rays with lowest points in L, $\mu_1 \leqslant \lambda \leqslant 1$, and (35)–(40) apply.

For rays with lowest points in N, $\lambda \leqslant \mu_1$, and we deduce from (45)–(47) that

$$\Delta = \alpha_0\{\cos^{-1}\lambda - (1-k)\cos^{-1}(\lambda/\mu_1)\}, \tag{48}$$

$$T = \alpha_0\{(1-\lambda^2)^{\frac{1}{2}} - (1-k)(\mu^2-\lambda^2)^{\frac{1}{2}}\}, \tag{49}$$

$$d\Delta/d\lambda = \alpha_0\{-(1-\lambda^2)^{-\frac{1}{2}} + (1-k)(\mu_1^2-\lambda^2)^{-\frac{1}{2}}\}. \tag{50}$$

By (37), (38), (48) and (49), Δ and T are thus continuous functions of λ for the family of all rays confined to L and N.

On the other hand, at $\lambda = \mu_1$, (36) and (50) show that $d\Delta/d\lambda$ jumps discontinuously (as λ decreases) from a finite negative value to $+\infty$, and then decreases through finite positive values. The behaviour of $dT/d\lambda$ is similar since $dT/d\lambda = \lambda\, d\Delta/d\lambda$. Thus Δ and T, as functions of λ, have node-point maxima at $\lambda = \mu_1$. Since $dT/d\Delta$, i.e. λ, is continuously varying, it follows that there is a cusp at the corresponding point A (Figure 1) of the graph of T against Δ; the branch OA corresponds to $1 \geqslant \lambda \geqslant \mu_1$.

As λ decreases from the value μ_1, the T–Δ curve recedes from A along the upper branch shown in Figure 1. Provided N is deep enough (50) shows that Δ and T

must pass, this time with continuous λ-gradients, through minima at which λ satisfies

$$\mu_1^2-\lambda^2 = (1-k)^2(1-\lambda^2); \tag{51}$$

correspondingly, the T–Δ curve passes through a second cusp B (Figure 1) and then advances, crossing OA at C.

The equations enable the points A, B and C to be precisely located when α_0, μ_1 and k are known. At the cusp A, $\Delta = \alpha_0 \cos^{-1}\mu_1$ and $T = \alpha_0\sqrt{(1-\mu_1^2)}$. The cusp B is located using (51) and then (48) and (49). For the crossing-point C, the equations (38), (39), (48) and (49) yield Δ, T and the two values of λ.

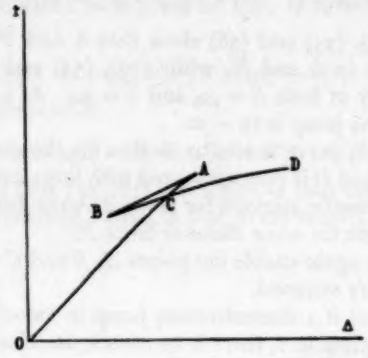


FIG. 1.

If the velocity distribution is such that the bottom of N is reached before the value of λ given by (51) is attained, then the curve of Figure 1 is relevant only up to some point K on the branch AB. The form taken after K will depend on the velocity distribution below N.

The present analysis, which is more definitive than that in the writer's text (1959), also yields the following results:

(i) Where v is continuous, any discontinuous increase in its gradient even if small, is necessarily accompanied by triplication in the time-curve for rays terminating at r_0.

(ii) Although A is a cusp, it is not associated with large amplitudes since $d\Delta/d\lambda$, and therefore $d\Delta/de_0$, is not zero at A.

(iii) On the other hand, $d\Delta/de_0$ is zero at the cusp B, which is thus associated with a caustic.

(iv) On the branch AB, $d\lambda/d\Delta$ and therefore $d^2T/d\Delta^2$ is zero at A. Hence the branch runs virtually straight for some distance from A towards B, any marked curvature being near B.

5.2 Discontinuous increase in velocity

We next consider the case $v_b > v_a$, so that $\mu_b < \mu_a$, other circumstances being as in Section 5.1.

Rays for which $\mu_a \leqslant \lambda \leqslant 1$ lie entirely in L, and (35)–(40) apply as before.

Rays for which $\mu_b \leqslant \lambda < \mu_a$ also lie entirely in L, but consist of two segments with an upward reflection point at the boundary r_1. For these rays, (45)–(47) give

$$\Delta = \alpha_0\{\cos^{-1}\lambda - \cos^{-1}(\lambda/\mu_a)\}, \tag{52}$$

$$T = \alpha_0\{(1-\lambda^2)^{\frac{1}{2}} - (\mu_a - \lambda^2)^{\frac{1}{2}}\}, \tag{53}$$

$$d\Delta/d\lambda = \alpha_0\{-(1-\lambda^2)^{-\frac{1}{2}} + (\mu_a{}^2 - \lambda^2)^{-\frac{1}{2}}\}. \tag{54}$$

For rays whose lowest points are in N, $\lambda < \mu_b$; and

$$\Delta = \alpha_0\{\cos^{-1}\lambda - \cos^{-1}(\lambda/\mu_a) + k\cos^{-1}(\lambda/\mu_b)\}, \tag{55}$$

$$T = \alpha_0\{(1-\lambda^2)^{\frac{1}{2}} - (\mu_a{}^2 - \lambda^2)^{\frac{1}{2}} + k(\mu_b{}^2 - \lambda^2)^{\frac{1}{2}}\}, \tag{56}$$

$$d\Delta/d\lambda = \alpha_0\{-(1-\lambda^2)^{-\frac{1}{2}} + (\mu_a{}^2 - \lambda^2)^{-\frac{1}{2}} - k(\mu_b{}^2 - \lambda^2)^{-\frac{1}{2}}\}. \tag{57}$$

(37), (38), (52), (53), (55) and (56) show that Δ and T are again continuous functions of λ for rays in L and N, while (39), (54) and (57) show that $d\Delta/d\lambda$ changes discontinuously at both $\lambda = \mu_a$ and $\lambda = \mu_b$. At μ_a the jump is to $+\infty$ as before, while at μ_b the jump is to $-\infty$.

The form of the T–Δ curve is similar to that for the case of Section 5.1. But neither of the cusps A and B is now associated with large amplitudes. The branch AB, as before, runs virtually straight for some distance from A, and the branch BD likewise runs straight for some distance from B.

The given equations again enable the points A, B and C to be precisely located when α_0, k, μ_a and μ_b are assigned.

It is to be noted that if a discontinuous jump in velocity gradient is accompanied by even a small jump in v, there is no caustic associated with either A or B.

5.3 Discontinuous decrease in velocity

When $v_b < v_a$, we have $\mu_b > \mu_a$. In treating this case, we assume that N extends downward far enough for μ to return through the value μ_a before the bottom of N is reached. (The analysis can be adapted to meet the additional complexity when this assumption does not hold.)

For rays with lowest points in L, (35)–(40) still apply. For rays which penetrate N, (55)–(57) apply; but since $\lambda < \mu_a$ for these rays, there are no rays with lowest points in the upper part of N, for which $\mu_a < \lambda < \mu_b$.

For the whole family of rays, there are thus, at $\lambda = \mu_a$, discontinuities $\delta\Delta$ and δT in Δ and T, given by

$$\delta\Delta = k\alpha_0\cos^{-1}(\mu_a/\mu_b), \tag{58}$$

$$\delta T = k\alpha_0(\mu_b{}^2 - \mu_a{}^2)^{-\frac{1}{2}}. \tag{59}$$

For the rays penetrating N, $d\Delta/d\lambda$ has, by (57), the value $+\infty$ at $\lambda = \mu_a$, and decreases as λ decreases from μ_a. Thus Δ, and also T, decrease rapidly as λ decreases from μ_a and (provided N is deep enough) reach minima given by the value of λ satisfying

$$(1-\lambda^2)^{-\frac{1}{2}} + k(\mu_b{}^2 - \lambda^2)^{-\frac{1}{2}} = (\mu_a{}^2 - \lambda^2)^{-\frac{1}{2}}. \tag{60}$$

The T–Δ graph is therefore as in Figure 2, the cusp at C corresponding to a caustic in this case. Near B, the branch BC is virtually straight.

The equations (58)–(60) enable B and C to be precisely located.

5.4 Rapid but continuous increase in velocity gradient

We now consider circumstances as in Section 5.1 except that the increase in velocity gradient from L to N takes place continuously, through a layer M for which $r_2 \leqslant r \leqslant r_1$, instead of abruptly at $r = r_1$; the upper boundary of N is now taken as $r = r_2$.

As previously, we take α equal to α_0, $k\alpha_0$ in L, N, respectively. For mathematical convenience we take

$$d\alpha/d\mu = c\mu \tag{61}$$

in M, where c is a positive constant, given by

$$2\alpha_0(1-k) = c(\mu_1^2-\mu_2^2). \tag{62}$$

For rays confined to L, $d\Delta/d\lambda$ is given by (35). For rays with lowest points in M, we have by (24) and (61)

$$d\Delta/d\lambda = -\alpha_0(1-\lambda^2)^{-\frac{1}{2}}+c(\mu_1^2-\lambda^2)^{\frac{1}{2}}. \tag{63}$$

For rays with lowest points in N,

$$d\Delta/d\lambda = -\alpha_0(1-\lambda^2)^{-\frac{1}{2}}+c(\mu_1^2-\lambda^2)^{\frac{1}{2}}-c(\mu_2^2-\lambda^2)^{\frac{1}{2}}. \tag{64}$$

It is seen from (35), (63) and (64) that $d\Delta/d\lambda$ is now continuous for the whole range of values of λ concerned. Also, Δ and T are readily seen to be continuous.

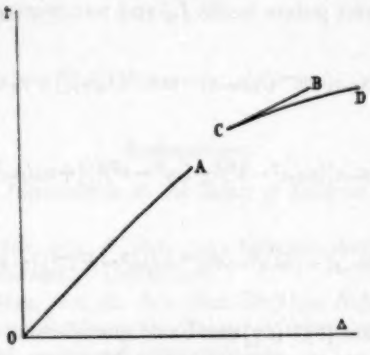


FIG. 2.

It is shown elsewhere (1960a) that, for a wide range of values of μ_1 and μ_2 that can be of importance to conditions in the Earth, $d\Delta/d\lambda$ passes through two zeros, changing from negative to positive and then back to negative, as λ decreases continuously from its initial value of unity.

Hence there can be triplication in the T–Δ curve again, but in this case large amplitudes will be associated with both the points A and B of Figure 1.

5.5 Remarks

The cases of Sections 5.2 and 5.3 are commonly said to involve "first-order" discontinuities in v, and that of Section 5.1 a "second-order" discontinuity. The case of Section 5.4 would then formally involve a "third-order" discontinuity. Our argument shows that a third- or even higher-order discontinuity can have an important influence on wave amplitudes. It follows that attempts to infer features

of velocity distribution from amplitude observations must be treated with great caution.

The various cases here treated are of course mathematical models, and assume constancy of α in each of the regions L and N. Previous work shows, however, that this assumption is extremely satisfactory over limited ranges of depth where the behaviour of v is normal. Thus the theory may be applied quantitatively, provided r_0 is taken as the radius of a suitably "stripped" Earth.

An important feature of the models is that they show clearly the circumstances in which cusps in the travel-time curves are associated with abnormally large amplitudes. Attention has already been drawn (1960a) to the error of assuming that cusps and large amplitudes necessarily go together.

6. Multi-layering

The cases of Sections 5.1 to 5.3 involved a single discontinuity surface below the outer surface $r = r_0$. The analysis can be readily extended to deal with cases of multi-layering.

Take for example the case where there is a series of layers L_0, L_1, L_2, ..., bounded above by $r = r_0, r_1, r_2, \ldots$, respectively. Let α take the constant values $\alpha_0, \alpha_1, \alpha_2, \ldots$, in the respective layers, and let v be continuous throughout. Then if the suffix n applies to the level r_n, we easily deduce from (35)–(39) and (45)–(47) that, for rays with lowest points inside L_n and terminating at r_0,

$$\Delta = \sum_{r=1}^{n} [\alpha_{r-1}\{\cos^{-1}(\lambda/\mu_{r-1}) - \cos^{-1}(\lambda/\mu_r)\}] + \alpha_n \cos^{-1}(\lambda/\mu_n), \tag{65}$$

$$T = \sum_{r=1}^{n} [\alpha_{r-1}\{(\mu_{r-1}^2 - \lambda^2)^{\frac{1}{2}} - (\mu_r^2 - \lambda^2)^{\frac{1}{2}}\}] + \alpha_n(\mu_n^2 - \lambda^2)^{\frac{1}{2}}, \tag{66}$$

$$d\Delta/d\lambda = \sum_{r=1}^{n} [\alpha_{r-1}\{-(\lambda_{r-1}^2 - \lambda^2)^{-\frac{1}{2}} + (\mu_r^2 - \lambda^2)^{-\frac{1}{2}}\}] - \alpha_n(\mu_n^2 - \lambda^2)^{-\frac{1}{2}}. \tag{67}$$

Formulae of the form (65)–(67) can be of special value when modern computing methods are used to set up travel-time tables from model velocity distributions. A usual procedure in representing possible velocity variations in parts of the Earth has been to take models consisting of layers in each of which v is constant. By taking instead, as in (65)–(67), layers in each of which α is constant, greater flexibility is provided. Discontinuities in v are not introduced, and, further, a suitable representation of the velocity distribution over a wide range of depth may be contrived with far fewer layers.

It is easy to make the necessary modifications to (65)–(67) when discontinuities in v are postulated. On the basis of the analysis presented, a wide class of problems can in fact be examined in quantitative detail.

7. Approximation for small Δ

Correct to the third order in Δ, the formula (39), which holds when α is constant, becomes

$$T = \Delta - (6\alpha_0^2)^{-1}\Delta^3. \tag{68}$$

In practice, a cubic form equivalent to (68) has often been used to represent T–Δ data for small or moderate Δ in a suitably stripped Earth.

When (68) applies, we have by (17)

$$\lambda = 1-\Delta^2/2\alpha_0^2; \quad d\Delta/d\lambda = -\alpha_0^2/\Delta. \tag{69}$$

Then, by (28) and (69),

$$\begin{aligned}\theta &= -\alpha_0^2\Delta^{-1}+\alpha_0\{1-(1-\Delta^2/2\alpha_0^2)^2\}^{-\frac{1}{2}} \\ &\approx \tfrac{1}{8}\Delta(1+3\Delta^2/16\alpha_0^2).\end{aligned} \tag{70}$$

Also, since $\psi = \lambda\theta$, we have by (69) and (70)

$$\psi \approx \tfrac{1}{8}\Delta(1-5\Delta^2/16\alpha_0^2). \tag{71}$$

With the use of (70) or (71), it is possible to derive the velocity variation corresponding to (68) very rapidly; substitution of (70) into (33), or (71) into (34) gives the value of $\alpha-\alpha_0$ for any assigned μ. In practice, it is sufficient to take

$$\theta \approx \psi \approx \tfrac{1}{8}\Delta, \tag{72}$$

in place of (69) and (70), for rays with Δ up to an appreciable fraction of a radian, and to carry out a simple numerical integration of (33) or (34). The process is illustrated in an earlier paper (1960).

University of Sydney:

1960 *June* 29.

References

Bôcher, M., 1914. *An Introduction to the Study of Integral Equations* (Cambridge: University Press).

Bullen, K. E., 1945. *Mon. Not. R. Astr. Soc. Geophys. Suppl.*, **5**, 91.

Bullen, K. E., 1954. *Seismology* (Methuen).

Bullen, K. E., 1955. *Mon. Not. R. Astr. Soc. Geophys. Suppl.*, **7**, 49.

Bullen, K. E., 1959. *Introduction to the Theory of Seismology* (Cambridge: University Press), 2nd ed., reprinted, with revisions.

Bullen, K. E., 1960. *Geophys. J.* **3**, 258.

Bullen, K. E., 1960a. *Geophys. J.* **3**, 354.

Jeffreys, Sir Harold, 1939. *Mon. Not. R. Astr. Soc. Geophys. Suppl.*, **4**, 508.

Jeffreys, Sir Harold, 1959. *Geophys. J.* **1**, 191.

Lehmann, I., 1959. *Ann. de Géophys.*, **15**, 93.

Polarization of Transverse Seismic Waves

Markus Båth

"And perpendicular now and now transverse,
Pierce the dark soil and as they pierce and pass
Make bare the secrets of the Earth's deep heart."

P. B. Shelley, Prometheus Unbound.

Summary

A theoretical investigation is made of the changes of the polarization of transverse seismic waves during their propagation through the Earth. The polarizations have been computed theoretically and numerically for reflexion at the core boundary and at the Earth's surface, for refraction and reflexion at the base of the crust and for passages through continuously varying media. It is demonstrated that great changes of the vibration properties (vibration angle and particle orbit) may occur in all cases except for continuously varying media, through which transverse waves propagate with practically unchanged vibration properties. The consequences of these results for earthquake mechanism studies, based on transverse waves, are discussed.

1. Introduction

Transverse seismic waves or *S* waves are composed of two components, *SH* and *SV*. In studies of *S*-wave propagation and especially of the behaviour at discontinuity surfaces it is most convenient to investigate *SH* and *SV* separately. The equations governing reflexion and refraction of *S* waves are of such a nature that a separate treatment of *SH* and *SV* is possible. But the actual particle motion is the result of *SH* and *SV* motions, and a combination of *SH* and *SV* is necessary in studies of the *S*-wave polarization.

In this paper we shall present a theoretical investigation of *S*-wave polarization and especially its changes upon reflexion or refraction at discontinuity surfaces in the Earth as well as during passages through continuously varying media. The results may be of interest in elucidating some wave propagation phenomena and particularly with regard to earthquake mechanism studies, using *S* waves.

2. Notation

Reflexion and refraction of seismic waves are dealt with in most textbooks in seismology, and therefore the fundamental ideas need only be briefly mentioned here. In our investigation of reflexion and refraction of *S* waves we shall apply

the method used by Jeffreys (1926), and our notation will be almost the same as used by Jeffreys:

x, y, z = rectangular coordinates; the waves propagate in the xz plane and $z = 0$ coincides with a discontinuity surface;
u, v, w = displacements along x, y, z respectively;
Φ, Ψ = potentials, such that

$$u = \frac{\partial \Phi}{\partial x} - \frac{\partial \Psi}{\partial z} \quad \text{and} \quad w = \frac{\partial \Phi}{\partial z} + \frac{\partial \Psi}{\partial x}; \tag{1}$$

λ, μ = Lamé's parameters;
ρ = density;
c_P, c_S = wave velocities for P, S respectively;
t = time;

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial z^2};$$

p_{zz}, p_{zx}, p_{zy} = normal and tangential stresses acting on a discontinuity surface ($z = 0$);
A, B, C = amplitude functions for P, SV, SH respectively;
A_{SH}, A_{SV} = amplitudes of SH, SV respectively;
e = angle of emergence
δ = vibration angle (see Section 3);
Δ = epicentral distance.

Quantities related to incident, reflected and transmitted (refracted) waves are denoted as follows:

incident waves: no accent or suffix (B, C, etc.),
reflected waves: suffix unity (B_1, C_1, etc.),
transmitted waves: accent (B', C', etc.).

Similarly, quantities related to the incident side of a discontinuity are given with no accent (c_P, c_S, ρ, etc.), those belonging to the other side have an accent (c_P', c_S', ρ', etc.).

Any part of Φ, Ψ, v can be written as follows with usual notation:

$$\left.\begin{aligned} \Phi &\sim A\exp[i\kappa(\alpha z + x - \omega t)] \\ \Psi &\sim B\exp[i\kappa(\beta z + x - \omega t)] \\ v &\sim C\exp[i\kappa(\gamma z + x - \omega t)]. \end{aligned}\right\} \tag{2}$$

Φ, Ψ, v satisfy the equations of motion:

$$\left.\begin{aligned} \rho\frac{\partial^2 \Phi}{\partial t^2} &= (\lambda + 2\mu)\nabla^2 \Phi \\ \rho\frac{\partial^2 \Psi}{\partial t^2} &= \mu\nabla^2 \Psi \\ \rho\frac{\partial^2 v}{\partial t^2} &= \mu\nabla^2 v. \end{aligned}\right\} \tag{3}$$

Inserting the expressions (2) of Φ, Ψ, v into (3) we find

$$\begin{aligned}\omega^2 &= c_P^2(1+\alpha^2) = c_S^2(1+\beta^2) = c_S^2(1+\gamma^2) \\ &= c_P'^2(1+\alpha'^2) = c_S'^2(1+\beta'^2) = c_S'^2(1+\gamma'^2).\end{aligned} \tag{4}$$

Furthermore, $\alpha = \tan e$ (P wave), $\beta = \tan e$ (SV), $\gamma = \tan e$ (SH). The equations of motion evidently express only Snell's refraction law. In case $\lambda = \mu$ (Poisson's relation) holds, i.e. $c_P = c_S\sqrt{3}$, we get $\beta^2 = 2+3\alpha^2$.

Those stress components, which will be used, have the following expressions:

$$\left.\begin{aligned} p_{zz} &= \lambda\nabla^2\Phi + 2\mu\left(\frac{\partial^2\Phi}{\partial z^2} + \frac{\partial^2\Psi}{\partial z\,\partial x}\right) \\ p_{zx} &= \mu\left(2\frac{\partial^2\Phi}{\partial z\,\partial x} + \frac{\partial^2\Psi}{\partial x^2} - \frac{\partial^2\Psi}{\partial z^2}\right) \\ p_{zy} &= \mu\frac{\partial v}{\partial z}. \end{aligned}\right\} \tag{5}$$

Our calculations are valid only for plane wave fronts and for plane discontinuity surfaces. The results are therefore not valid near an earthquake focus, where the front cannot be assumed plane.

3. Determination of the vibration angle

The vibration plane is defined as a plane along the direction of wave propagation and oriented such that the vibration takes place in this plane. The vibration angle δ is the angle between the vibration plane and the vertical plane of propagation. Therefore, looking in the direction of propagation, we have for linearly polarized S waves

$$\tan\delta = \frac{A_{SH}}{A_{SV}}. \tag{6}$$

δ is counted clockwise from the vertical direction from 0° to 180° ($0° \leqslant \delta \leqslant 180°$).

In all our computations we assume the incident S wave to be linearly polarized, whereas the reflected or refracted S wave may be linearly or elliptically polarized, depending upon the angle of emergence. An elliptically polarized S wave corresponds to a complex expression for B_1/B or B'/B.

For an incident S wave we have

$$\left.\begin{aligned} SH:\ v &= C\exp[i\kappa(x+\gamma z-\omega t)] \\ SV:\ \Psi &= B\exp[i\kappa(x+\beta z-\omega t)]. \end{aligned}\right\} \tag{7}$$

The expressions for v and Ψ give us the amplitudes in the incident wave:

$$\left.\begin{aligned} A_{SH} &= C \\ A_{SV} &= i\kappa B(1+\beta^2)^{\frac{1}{2}} \end{aligned}\right\} \tag{8}$$

remembering that A_{SV} is the maximum amplitude of $(u^2+w^2)^{\frac{1}{2}}$. We thus find

$$\tan\delta = \frac{C}{i\kappa B(1+\beta^2)^{\frac{1}{2}}}. \tag{9}$$

Similarly, we find for a reflected S wave:

$$\left.\begin{aligned} v_1 &= C_1 \exp[i\kappa(x-\gamma z-\omega t)] \\ \Psi_1 &= B_1 \exp[i\kappa(x-\beta z-\omega t)] \end{aligned}\right\} \tag{10}$$

which give

$$\tan\delta_1 = \frac{A_{SH1}}{A_{SV1}} = \frac{C_1}{i\kappa B_1(1+\beta^2)^{\frac{1}{2}}} \tag{11}$$

and

$$\frac{\tan\delta_1}{\tan\delta} = \frac{C_1/C}{B_1/B}. \tag{12}$$

For a transmitted (refracted) S wave we get in the same way:

$$\left.\begin{aligned} v' &= C' \exp[i\kappa(x+\gamma' z-\omega t)] \\ \Psi'' &= B' \exp[i\kappa(x+\beta' z-\omega t)] \end{aligned}\right\} \tag{13}$$

$$\tan\delta' = \frac{A_{SH}'}{A_{SV}'} = \frac{C'}{i\kappa B'(1+\beta'^2)^{\frac{1}{2}}} \tag{14}$$

and

$$\frac{\tan\delta'}{\tan\delta} = \frac{(1+\beta^2)^{\frac{1}{2}}}{(1+\beta'^2)^{\frac{1}{2}}}\frac{C'/C}{B'/B} = \frac{c'_S}{c_S}\frac{C'/C}{B'/B}. \tag{15}$$

These formulae can be used for computation of the vibration angle only for linearly polarized waves. In case of elliptical polarization, i.e. B_1/B or B'/B being complex quantities, we let the vibration plane coincide with the major axis of the ellipse. New formulae are then required for the vibration angle.

For a reflected, elliptically polarized S wave we may put

$$\frac{A_{SH1}}{A_{SH}} = \frac{C_1}{C} = c; \qquad \frac{A_{SV1}}{A_{SV}} = \frac{B_1}{B} = a+ib. \tag{16}$$

The complex expression for the reflected SV wave corresponds to an amplitude $= (a^2+b^2)^{\frac{1}{2}}$ and a phase shift $= \tan^{-1}(b/a)$ compared to the incident SV. The ellipse, representing the particle motion, has the following equation, ξ and η being rectangular coordinates in the plane of the ellipse with ξ horizontal:

$$\frac{(a^2+b^2)A_{SV}^2}{c^2A_{SH}^2}\xi^2+\eta^2-2\frac{aA_{SV}}{cA_{SH}}\xi\eta = b^2A_{SV}^2. \tag{17}$$

We rotate the $\xi\eta$ axes so as to coincide with the axes of the ellipse, which leads to the following formula for δ_1:

$$\tan 2\delta_1 = \frac{2ac\tan\delta}{a^2+b^2-c^2\tan^2\delta}. \tag{18}$$

Similarly, in case of an elliptically polarized, transmitted wave:

$$\frac{A_{SH}'}{A_{SH}} = \frac{C'}{C} = c; \qquad \frac{A_{SV}'}{A_{SV}} = \frac{c_S}{c_S'}\frac{B'}{B} = a+ib \tag{19}$$

$$\tan 2\delta' = \frac{2ac\tan\delta}{a^2+b^2-c^2\tan^2\delta}. \tag{20}$$

Note the meaning of a, b in this case as compared to the case of reflected waves. The special case of circular polarization will arise if the conditions $a = 0$ and $\tan\delta = b/c$ are both fulfilled; the vibration angle (δ_1 or δ') is then indefinite.

The formulae (18) and (20) reduce to the corresponding equations (12) and (15) for linearly polarized waves, if we put $b = 0$. The equations (18) and (20) are more general, and they can be used in all cases, when the incident vibration is linear.

4. Reflexion at the Earth's core (*ScS*)

For an *SV* wave incident upon the Earth's core we have the following expressions for the potentials, considering the fact that there is no transmitted *S* wave:

$$\left.\begin{aligned} \Psi &= B\exp[i\kappa(x-\beta z-\omega t)]+B_1\exp[i\kappa(x+\beta z-\omega t)] \\ \Psi' &= 0 \\ \Phi &= A_1\exp[i\kappa(x+\alpha z-\omega t)] \\ \Phi' &= A'\exp[i\kappa(x-\alpha' z-\omega t)]. \end{aligned}\right\} \tag{21}$$

We have assumed $\lambda \neq \mu$ in this case, which is also in agreement with our present knowledge of λ and μ at the core boundary, and we use the complete expressions (4), obtained from the equations of motion. The boundary conditions require continuity of w, p_{zz} and p_{zx} at $z = 0$, i.e. using equations (1) and (5):

$$\left.\begin{aligned} \frac{B_1}{B}+\frac{A_1}{B}\alpha+\frac{A'}{B}\alpha' &= -1 \\ \frac{B_1}{B}+\frac{A_1}{B}\frac{2\alpha}{1-\beta^2} &= -1 \\ \frac{B_1}{B}+\frac{A_1}{B}\frac{\alpha^2+1-2\left(\frac{c_S}{c_P}\right)^2}{2\beta\left(\frac{c_S}{c_P}\right)^2}-\frac{A'}{B}\frac{\frac{\rho'}{\rho}c_P'^2(1+\alpha'^2)}{2\beta c_S^2} &= 1. \end{aligned}\right\} \tag{22}$$

The unknowns are B_1/B, A_1/B, A'/B. Solving for B_1/B we find

$$\frac{B_1}{B} = -1+\frac{2}{1+\frac{\rho'(1+\beta^2)^2}{4\rho\alpha'\beta}+\frac{(1-\beta^2)^2}{4\alpha\beta}}. \tag{23}$$

For an *SH* wave incident at the core boundary the corresponding expressions are

$$\left.\begin{aligned} v &= C\exp[i\kappa(x-\gamma z-\omega t)]+C_1\exp[i\kappa(x+\gamma z-\omega t)] \\ v' &= 0. \end{aligned}\right\} \tag{24}$$

The boundary condition is that p_{zy} is continuous at $z = 0$, which gives

$$\mu(-\gamma C+\gamma C_1) = 0 \tag{25}$$

or $C_1/C = +1$ for all angles of emergence.

Numerical values are given in Table 1, computed for $\rho = 5{\cdot}4\,\mathrm{g/cm^3}$, $\rho' = 10{\cdot}1\,\mathrm{g/cm^3}$, $c_S = 7{\cdot}25\,\mathrm{km/s}$, $c_P = 13{\cdot}7\,\mathrm{km/s}$, $c_P' = 8{\cdot}0\,\mathrm{km/s}$. In addition to B_1/B we give numerical values of A_1/B and A'/B in order to be complete, even if the latter values are of no further use in this investigation. All numerical results have been

Table 1

Reflexion of SV at the Earth's outer core

e deg	B_1/B	A_1/B	A'/B
0	-1	0	0
10	$-0{\cdot}9744+0{\cdot}2248i$	$-0{\cdot}1292+0{\cdot}0147i$	$-0{\cdot}2962+0{\cdot}0337i$
20	$-0{\cdot}9583+0{\cdot}2857i$	$-0{\cdot}1485+0{\cdot}0217i$	$-0{\cdot}6118+0{\cdot}0892i$
30	$-0{\cdot}6478+0{\cdot}0151i$	$-0{\cdot}0064+0{\cdot}1483i$	$-0{\cdot}7614-0{\cdot}0327i$
40	$-0{\cdot}4447+0{\cdot}0056i$	$-0{\cdot}0011+0{\cdot}1136i$	$-0{\cdot}7606-0{\cdot}0076i$
50	$-0{\cdot}3965+0{\cdot}0119i$	$+0{\cdot}0044-0{\cdot}2235i$	$-0{\cdot}7349-0{\cdot}0145i$
60	$-0{\cdot}6378$	$+1{\cdot}0437$	$-0{\cdot}4792$
70	$-0{\cdot}7866$	$+0{\cdot}5917$	$-0{\cdot}3717$
80	$-0{\cdot}9410$	$+0{\cdot}3191$	$-0{\cdot}1947$
90	-1	0	0

checked by the energy equation, expressing the fact that the incident energy equals the sum of the energies of the reflected and transmitted waves:

$$\left(\frac{B_1}{B}\right)^2+\frac{\alpha}{\beta}\left(\frac{A_1}{B}\right)^2+\frac{\alpha'\rho'}{\beta\rho}\left(\frac{A'}{B}\right)^2=1. \tag{26}$$

Using the values of Table 1 it is possible to calculate δ_1 for any given values of δ and e by means of equation (18). The results are shown in Figure 1 for a representative selection of e-values. The curves for $e = 10°$, $20°$ and $80°$ (not shown in Figure 1) are close to the straight line for $e = 0°$, $90°$, and $e = 30°$ is close to the curve for $e = 60°$. The inset figure shows the relation between the angle of emergence (e) at the core and the epicentral distance (Δ) for an *ScS* wave from a surface focus, computed by means of Jeffreys–Bullen travel times (1940).

Figure 1 demonstrates that for all angles of emergence at the core, the vibration angles δ and δ_1 lie in different quadrants. This naturally means a considerable change of the particle motion due to the reflexion. But the reason is easy to visualize physically, as it is sufficient with a phase reversal of one of the *S* components to bring this about.

5. Reflexion at the Earth's surface (*sS*, *SS*, ...)

This case was investigated by Jeffreys (1926, pp. 328–329 for *SV* and p. 324 for *SH*). The results are as follows:

$$\frac{B_1}{B}=\frac{4\alpha\beta-(1+3\alpha^2)^2}{4\alpha\beta+(1+3\alpha^2)^2} \tag{27}$$

and

$$\frac{C_1}{C}=+1. \tag{28}$$

In this solution we have assumed $\lambda = \mu$ or $\beta^2 = 2+3\alpha^2$, which is fairly true near the surface. The energy equation for the *SV* wave reads as follows:

$$\left(\frac{B_1}{B}\right)^2 + \frac{\alpha}{\beta}\left(\frac{A_1}{B}\right)^2 = 1. \qquad (29)$$

Table 2 gives numerical values for a series of e-values, and Figure 2 shows the relation between δ_1 and δ. The inset figure $e(\Delta)$ refers to S waves from a surface focus, assuming $c_S = 3{\cdot}2$ km/s. Unless the wavelength is very short, this

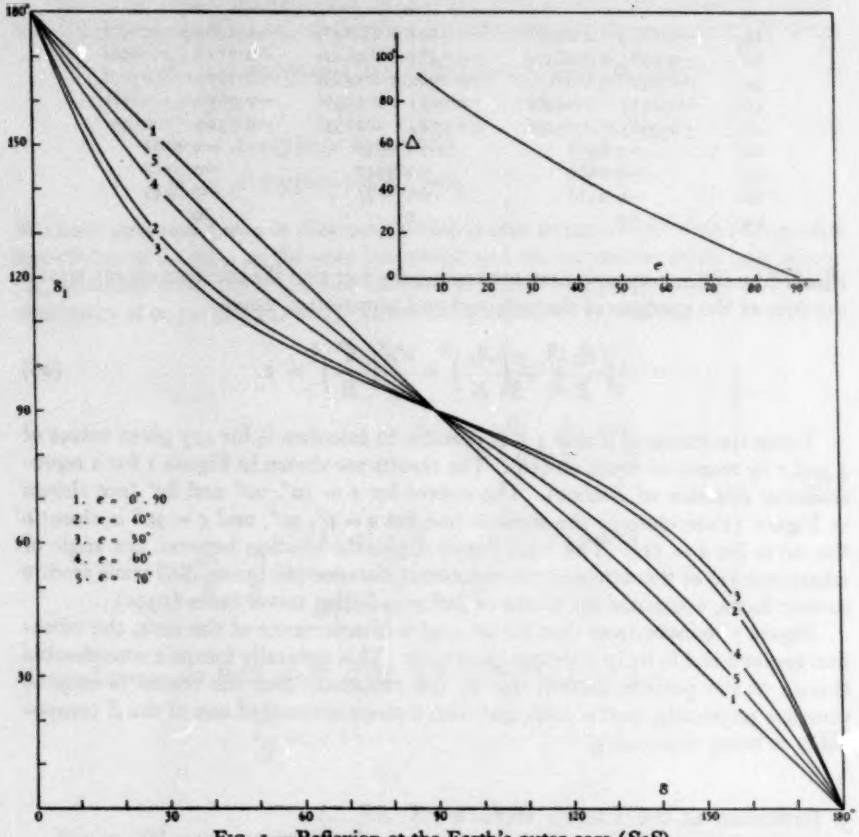


FIG. 1.—Reflexion at the Earth's outer core (*ScS*).

may be too low a value for c_S. The value immediately below Moho ($c_S = 4{\cdot}4$ km/s, see Figure 4) would be more appropriate in most cases. The δ_1–δ curves in Figure 2 are not influenced by the value of c_S. If an *SS* wave is observed at a station at distance Δ_1, the e-value at the reflexion point corresponds to the distance $\Delta_1/2$ in these graphs.

The δ_1–δ relation is more complicated for reflexions at the surface of the Earth than for reflexions at the core boundary. Especially within certain ranges of the

Table 2

Reflexion of SV at the Earth's surface

e deg	B_1/B	A_1/B
0	−1	0
10	−0·4596 + 0·8881i	+0·5312 − 0·3232i
20	+0·3994 + 0·9168i	+0·5040 − 0·7693i
30	+0·8750 + 0·4841i	+0·2165 − 0·8386i
40	+0·9968 + 0·0793i	+0·0178 − 0·4495i
45	+1	0
50	+0·9859 + 0·1675i	−0·0801 + 0·9498i
55·7	0	−2·5373
57·2	+0·0715	−2·0468
60	0	−1·7320
62·5	−0·1160	−1·5938
70	−0·4829	−1·2446
80	−0·8621	−0·6778
90	−1	0

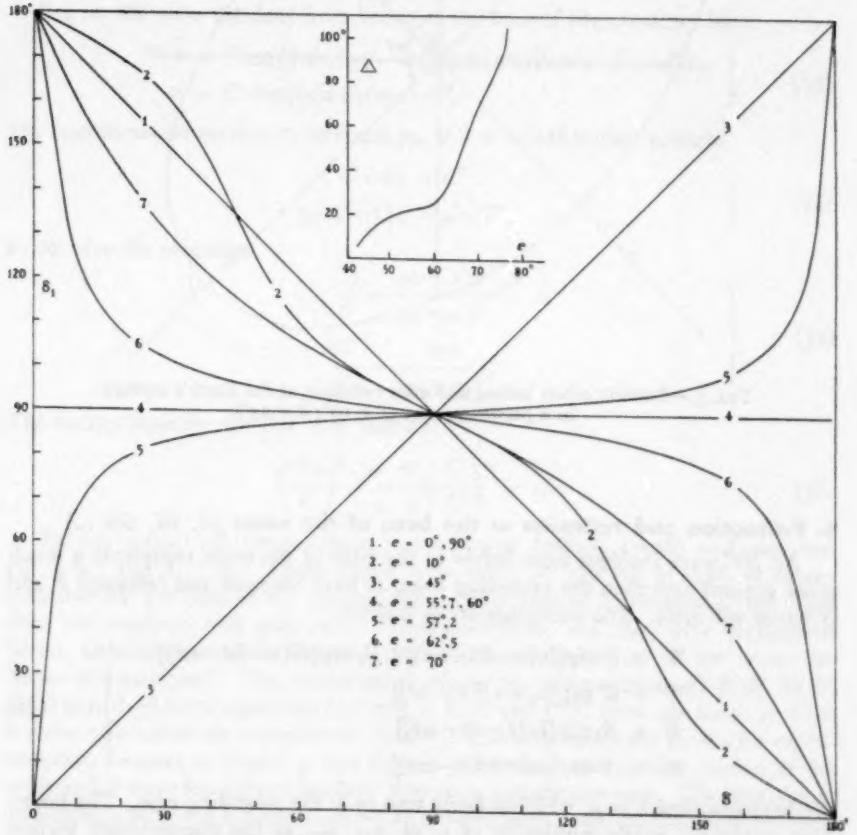


FIG. 2.—Reflexion at the Earth's surface (*sS*, *SS*,...).

e-values, there are rapid variations of the δ_1–δ relation for surface reflexions (Figure 2). Thus δ_1 and δ are in the same quadrant for $e = 15°{\cdot}0$–$54°{\cdot}2$ and for $e = 55°{\cdot}7$–$60°{\cdot}0$, but in different quadrants for other e-values. This is naturally only another way of expressing the conditions under which B_1 has the same or the opposite phase relative to B.

Figure 3 illustrates a particular case of reflexion at the Earth's surface, in which an incident linearly polarized wave gives rise to a reflected wave of elliptic polarization and with a different vibration angle.

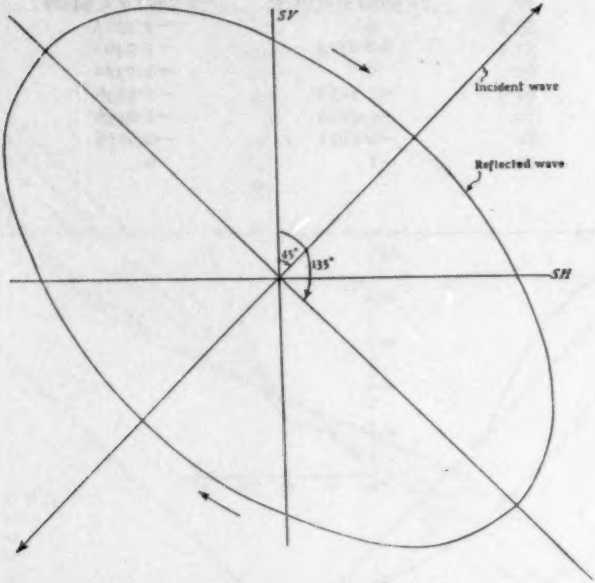


FIG. 3.—Particle orbits before and after reflexion at the Earth's surface in a particular case ($e = 10°$; $\delta = 45°$).

6. Refraction and reflexion at the base of the crust (S, sS, SS, ...)

An SV wave incident from below to the base of the crust represents a much more general case than the preceding ones, as both reflected and refracted P and S waves will arise. The potentials are as follows:

$$\left.\begin{aligned} \Psi &= B\exp[i\kappa(x+\beta z-\omega t)]+B_1\exp[i\kappa(x-\beta z-\omega t)] \\ \Psi' &= B'\exp[i\kappa(x+\beta' z-\omega t)] \\ \Phi &= A_1\exp[i\kappa(x-\alpha z-\omega t)] \\ \Phi' &= A'\exp[i\kappa(x+\alpha' z-\omega t)]. \end{aligned}\right\} \qquad (30)$$

We have assumed $\lambda = \mu$, which is fairly true as in the preceding case. The boundary conditions imply continuity of u, w, p_{zz}, p_{zx} at the discontinuity surface

($z = 0$), which gives the following equations:

$$\left.\begin{aligned} \frac{B'}{B}\beta' + \frac{B_1}{B}\beta - \frac{A'}{B} + \frac{A_1}{B} &= \beta \\ \frac{B'}{B} - \frac{B_1}{B} + \frac{A'}{B}\alpha' + \frac{A_1}{B}\alpha &= 1 \\ \frac{B'}{B}\frac{2\mu'\beta'}{\mu} + \frac{B_1}{B}2\beta + \frac{A'}{B}\frac{\mu'}{\mu}(1+3\alpha'^2) - \frac{A_1}{B}(1+3\alpha^2) &= 2\beta \\ \frac{B'}{B}\frac{\mu'}{\mu}(1-\beta'^2) - \frac{B_1}{B}(1-\beta^2) + \frac{A'}{B}\frac{2\mu'\alpha'}{\mu} + \frac{A_1}{B}2\alpha &= 1-\beta^2 \end{aligned}\right\} \quad (31)$$

with B'/B, B_1/B, A'/B, A_1/B as unknowns. The energy equation reads

$$\left(\frac{B_1}{B}\right)^2 + \frac{\beta'}{\beta}\frac{\rho'}{\rho}\left(\frac{B'}{B}\right)^2 + \frac{\alpha}{\beta}\left(\frac{A_1}{B}\right)^2 + \frac{\alpha'}{\beta}\frac{\rho'}{\rho}\left(\frac{A'}{B}\right)^2 = 1. \quad (32)$$

This is the general energy equation for *SV*, of which the earlier ones, (26) and (29), are immediately obtained as special cases.

For an *SH* wave incident from below at the base of the crust, we have

$$\left.\begin{aligned} v &= C\exp[i\kappa(x+\gamma z-\omega t)] + C_1\exp[i\kappa(x-\gamma z-\omega t)] \\ v' &= C'\exp[i\kappa(x+\gamma' z-\omega t)]. \end{aligned}\right\} \quad (33)$$

The conditions for continuity in v and p_{zy} at $z = 0$ lead to the relations

$$\left.\begin{aligned} C + C_1 &= C' \\ \mu\gamma(C - C_1) &= \mu'\gamma' C' \end{aligned}\right\} \quad (34)$$

which give the solutions

$$\left.\begin{aligned} \frac{C_1}{C} &= \frac{\mu\gamma - \mu'\gamma'}{\mu\gamma + \mu'\gamma'} \\ \frac{C'}{C} &= \frac{2\mu\gamma}{\mu\gamma + \mu'\gamma'}. \end{aligned}\right\} \quad (35)$$

The energy equation for *SH* is as follows:

$$\left(\frac{C_1}{C}\right)^2 + \frac{\mu'}{\mu}\frac{\gamma'}{\gamma}\left(\frac{C'}{C}\right)^2 = 1. \quad (36)$$

In our numerical computations we followed Jeffreys (1926) and assumed $c_S' = \frac{3}{4}c_S$ and $\rho' = \frac{4}{5}\rho$. These values were taken by Jeffreys in 1926 as representative for the base of the granitic layer. Our present knowledge of the crust does not conform well with these numerical values, and they may correspond better, although not perfectly, with the conditions at the base of the crust (the Moho discontinuity). The numerical results of the computations of B'/B, B_1/B, A'/B and A_1/B from equations (31) and of C_1/C and C'/C from equations (35) for a series of e-values are compiled in Table 3. The δ-relations are shown for representative e-values in Figure 4, and Figure 5 illustrates the particle motion of the refracted *S* wave for a given incident *S* wave in a particular case. The inset figure in Figure 4 gives the $e(\Delta)$-relation for an *S* wave from a surface focus and incident at Moho from below ($c_S = 4{\cdot}4$ km/s).

Table 3

Refraction and reflexion of SH and SV at Moho from below

ϵ	C'/C	C_1/C	B'/B	B_1/B	A'/B	A_1/B
deg						
0	0	−1	0	−1	—	—
10	+0·6007	−0·3993	+0·3350 + 0·1871*i*	−0·3100 + 0·5422*i*	—	—
20	+0·8911	−0·1089	+0·5663 + 0·1468*i*	+0·1300 + 0·4765*i*	—	—
30	+1·0458	+0·0458	+0·7139 + 0·0808*i*	+0·2736 + 0·2966*i*	—	—
40	+1·1338	+0·1338	+0·8115 + 0·0882*i*	+0·2095 + 0·1081*i*	+0·5689 − 0·3036*i*	—
50	+1·1861	+0·1861	+0·8226 + 0·0041*i*	+0·1975 + 0·0001*i*	+0·5535 − 0·0165*i*	—
60	+1·2179	+0·2179	+0·8985	+0·0579	+0·2859	−0·3035
70	+1·2367	+0·2367	+0·9221	−0·0971	+0·1798	−0·2803
80	+1·2530	+0·2530	+0·9415	−0·2032	+0·0864	−0·1634
90	+1·2500	+0·2500	+0·9375	−0·2500	0	0

There are reflected and refracted S waves for all angles of emergence, whereas a reflected P wave exists only for $e > 54°{\cdot}7$ and a refracted P wave only for $e > 39°{\cdot}7$. For the transmitted (refracted) S wave only one curve has been drawn in Figure 4 (for $e = 50°$), as the curves for all other e-values agree closely with the one given here. δ and δ' are in the same quadrant for all values of e. There are greater variations among the δ-curves for the reflected waves. δ and δ_1 lie in different quadrants for $e = 17°$–$25°$ and for $e > 64°$, but are in the same quadrant for other e-values. The δ-curve for the reflected S wave for $e = 50°$ agrees very closely with the curve for $e = 0°$ and has therefore been omitted from Figure 4.

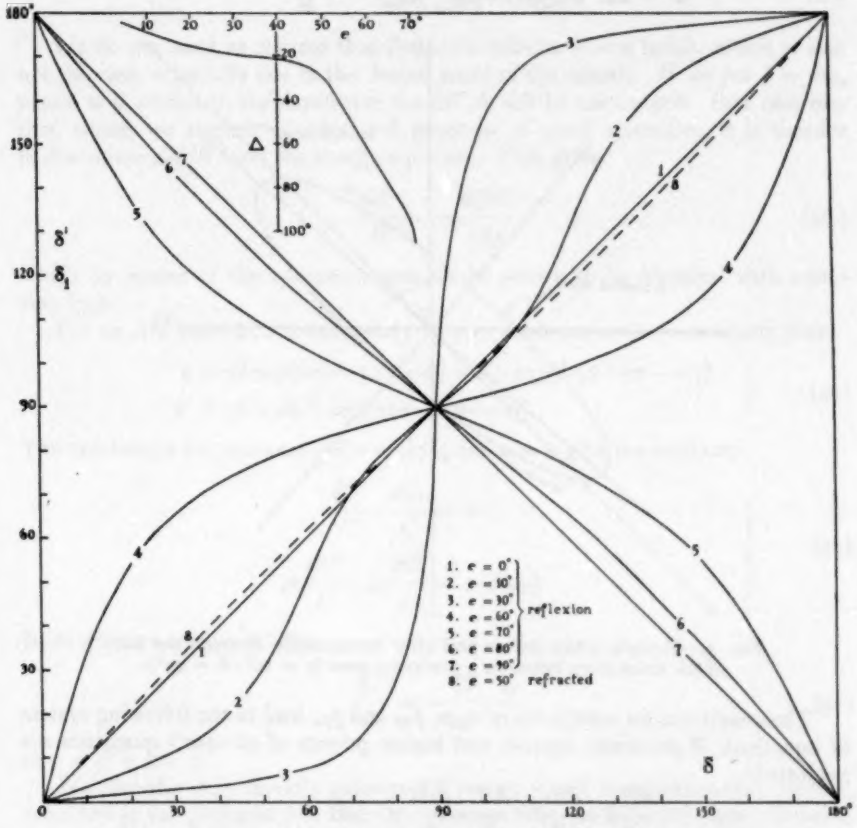


FIG. 4.—Refraction and reflexion at the base of the crust for incidence from below.

7. Transmission through continuously varying media (all transverse waves)

A continuously varying medium acts as if it were composed of a large number of thin layers with properties varying by small amounts from layer to layer, provided the wave length is short enough. Therefore, if a property belonging to

the incident side is denoted X, the corresponding property of the next layer is denoted $X' = X+dX$, e.g. $\rho' = \rho+d\rho$, $c_S' = c_S+dc_S$, $\beta' = \beta+d\beta$, $B' = B+dB'$, etc.

For an incident SV wave the potentials now become as follows, considering the fact that almost all of the incident energy passes into the transmitted wave of the same type, whereas all other waves will be represented by small quantities:

$$\left.\begin{aligned} \Psi &= B\exp[i\kappa(x-\beta z-\omega t)]+dB_1\exp[i\kappa(x+\beta z-\omega t)] \\ \Psi' &= (B+dB')\exp[i\kappa(x-\beta' z-\omega t)] \\ \Phi &= dA_1\exp[i\kappa(x+\alpha z-\omega t)] \\ \Phi' &= dA'\exp[i\kappa(x-\alpha' z-\omega t)]. \end{aligned}\right\} \tag{37}$$

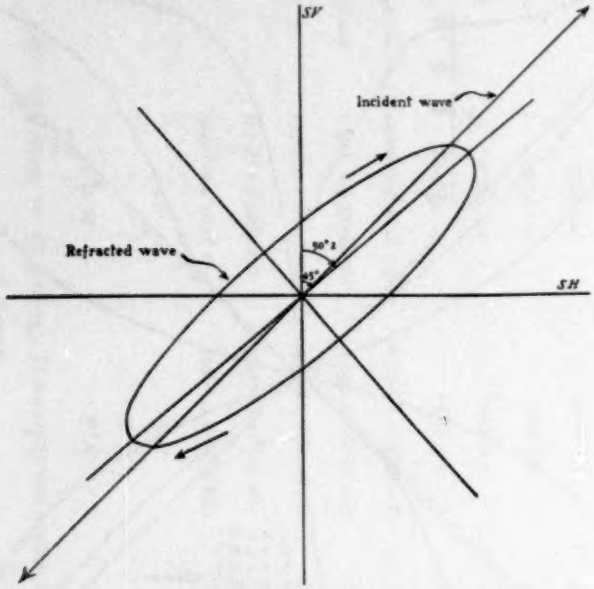


FIG. 5.—Particle orbits before and after transmission through the base of the crust from below in a particular case ($\epsilon = 10°$; $\delta = 45°$).

The conditions for continuity of u, w, p_{zz} and p_{zx} lead to the following system of equations, if products, squares and higher powers of all small quantities are neglected:

$$\left.\begin{aligned} \frac{dB'}{B}\beta+\frac{dB_1}{B}\beta+\frac{dA'}{B}-\frac{dA_1}{B} &= -d\beta \\ \frac{dB'}{B}-\frac{dB_1}{B}-\frac{dA'}{B}\alpha-\frac{dA_1}{B}\alpha &= 0 \\ \frac{dB'}{B}+\frac{dB_1}{B}-\frac{dA'}{B}p+\frac{dA_1}{B}p &= -\frac{d(\mu\beta)}{\mu\beta} \\ \frac{dB'}{B}-\frac{dB_1}{B}-\frac{dA'}{B}q-\frac{dA_1}{B}q &= -\frac{d[\mu(1-\beta^2)]}{\mu(1-\beta^2)} \end{aligned}\right\} \tag{38}$$

with

$$p = \frac{\lambda(1+\alpha^2)+2\mu\alpha^2}{2\mu\beta}$$

and

$$q = \frac{2\alpha}{1-\beta^2}.$$

The unknowns are dB'/B, dB_1/B, dA'/B, dA_1/B. Solving for dB'/B we find

$$\frac{dB'}{B} = -\frac{3\beta^2+1}{2\beta(1+\beta^2)}d\beta - \frac{d\mu}{2\mu}. \tag{39}$$

We do not need to assume that Poisson's relation $\lambda = \mu$ holds, which is also not the case, especially not in the deeper parts of the mantle. If we put $\lambda = m\mu$, where m is arbitrary, the expression for dB'/B will be unchanged. In a case like this, where we neglect squares and products of small quantities, it is simpler to determine dB'/B from the energy equation. This gives

$$\frac{dB'}{B} = -\frac{d(\beta\rho)}{2\beta\rho} \tag{39'}$$

which by means of the refraction law can be shown to be identical with equation (39).

For an *SH* wave in a continuously varying medium we have similarly that

$$\left.\begin{aligned} v &= C\exp[i\kappa(x-\gamma z-\omega t)]+dC_1\exp[i\kappa(x+\gamma z-\omega t)] \\ v' &= (C+dC')\exp[i\kappa(x-\gamma' z-\omega t)]. \end{aligned}\right\} \tag{40}$$

The conditions for continuity of v and p_{zy} for $z = 0$ give the relations

$$\left.\begin{aligned} \frac{dC'}{C}-\frac{dC_1}{C} &= 0 \\ \mu\gamma\frac{dC'}{C}+\mu\gamma\frac{dC_1}{C} &= -d(\mu\gamma) \end{aligned}\right\} \tag{41}$$

from which we solve dC'/C:

$$\frac{dC'}{C} = -\frac{d\beta}{2\beta}-\frac{d\mu}{2\mu} \tag{42}$$

as $\gamma = \beta = \tan e$.

We consider only linearly polarized S waves, which is approximately correct, considering the probable fact that the variation from an incident linear vibration will be quite small. Moreover, it is easy to demonstrate by examples that assuming elliptically polarized S waves as being linearly polarized, i.e. neglecting the imaginary term, is permitted as a good approximation in most computations of δ. Therefore, as in (15),

$$\frac{\tan\delta'}{\tan\delta} = \frac{A_{SH}'/A_{SH}}{A_{SV}'/A_{SV}} = \frac{C'/C}{B'(1+\beta'^2)^{\frac{1}{2}}/B(1+\beta^2)^{\frac{1}{2}}} = \frac{C'/C}{E'/E} \tag{43}$$

where $E = B(1+\beta^2)^{\frac{1}{2}}$.

For an infinitesimal variation we have $C' = C+dC'$, $E' = E+dE'$ and

$$\frac{\tan\delta'}{\tan\delta} = 1+\frac{dC'}{C}-\frac{dE'}{E}. \tag{44}$$

By integration of $[(dC'/C)-(dE'/E)]$ over a finite depth interval we should find the corresponding change in δ. Now

$$\frac{dE'}{E} = -\frac{d\beta}{2\beta}-\frac{d\mu}{2\mu} \tag{45}$$

and therefore

$$\frac{dC'}{C}-\frac{dE'}{E} = 0 \tag{46}$$

and

$$\delta' = \delta. \tag{47}$$

At the deepest point of the ray path, where total reflexion takes place, equations (37) and (38) are no longer valid. But it may be demonstrated in an analogous way or simpler by means of the energy equation that there is no change of δ in the deepest point either.

The result is that there is no change of the vibration angle in a continuously varying medium. This is valid under the assumptions we have made, i.e. in addition to those underlying all computations in this paper:

(1) that products, squares and higher powers of all small quantities may be neglected;

(2) that the waves may be considered as linearly polarized.

However, none of these assumptions will severely limit the validity of our result, which may therefore be considered at least as approximately true for the real Earth. This agrees with the empirical results of Monakhov (1950).

8. Reciprocity of the vibration angle

A question of importance, especially in connection with earthquake mechanism studies, concerns the possible reciprocity of the vibration angle. An incident S wave at a discontinuity surface has the vibration angle δ and the transmitted wave has the vibration angle δ'. If now an S wave with a vibration angle δ' is sent in exactly opposite direction, giving a transmitted S with vibration angle δ'', the question is if $\delta'' = \delta$ (exact reciprocity) or not.

As before, we consider only linearly polarized incident waves. We consider first transmission through Moho and will use the same system of equations for SV and SH as in Section 6, but now for incidence from above in addition to incidence from below. The details need not be given here. It may be shown that in case of linear polarization of the transmitted wave from below, the reciprocity is exact, i.e. $\delta'' = \delta$. But when the wave transmitted from below is elliptically polarized, the reciprocity breaks down and the wave transmitted back into the lower medium is in general also elliptically polarized. In a special case with $e = 20°{\cdot}0$ (lower

medium), $e' = 45°{\cdot}2$ (upper medium) we found the following corresponding values of δ and δ'':

δ	δ''
°	°
20	23·5
40	48·0
60	68·1
80	83·2

The condition for reciprocity of δ upon reflexion at any surface is that

$$\left|\frac{B_1}{B}\right| = \left|\frac{C_1}{C}\right|. \tag{48}$$

There is no reciprocity in δ for reflexions at Moho from below, except for $e = 0°$, 90°, as is evident from Table 3.

The vibration angle is also generally not reciprocal for reflexions at the Earth's core or at the Earth's surface, not even when the reflected wave is linearly polarized. The *SV* component is generally decreased at every reflexion, whereas *SH* is unchanged, i.e. δ increases for every reflexion. This corresponds to the observation that *S* waves reflected several times at the Earth's surface (*SSS*, ...) contain proportionally more of *SH* motion than of *SV*. Reciprocity occurs only for $e = 0°$, 90° for reflexions at the Earth's core (Table 1) and for $e = 0°$, 45°, 90° for reflexions at the Earth's surface (Table 2).

One consequence of these results is the following. If from observations of *S* waves at a seismograph station one wants to deduce the motion at the source, it is generally not permitted to follow the rays backwards to the source and correct for the changes on the way. Instead, it is necessary in making such corrections to follow the rays in the same direction as they propagate. Such a computation will require knowledge of reflexion and transmission coefficients at each interface.

9. Consequences for earthquake mechanism studies

The present study was begun already in 1946, when I had arrived at the idea that in earthquake mechanism studies use could be made of all seismic waves and not only of the first motion of *P* waves, which was practically the only procedure used at that time. I collected quite a number of readings of all possible phases both at Kew and at Uppsala, but these observations were not good enough to permit definite conclusions. I found it necessary to perform much theoretical computation, before the observations could possibly be interpreted. In 1946, I also visited Sir Harold Jeffreys in Cambridge for the first time and had the fortunate opportunity to discuss some of these problems with him. He drew my attention to his paper of 1926, which in fact has stimulated both the computations presented in this paper as well as a number of similar computations related to other waves.

Some investigations of the polarization of *S* waves were made already in the early days of instrumental seismology, as e.g. by Galitzin (1911) and by Geiger & Gutenberg (1912). Later Neumann (1930) published a paper on observations of *S* waves and their use in focal mechanism studies. In recent years many seismologists have used *S* waves for earthquake mechanism studies, e.g. Heinrich & Haill (1952), Dehlinger (1952), Ingram (1953), Gutenberg (1955), Keylis-Borok (1957),

Ritsema (1957), Adams (1958), Nuttli (1958), Byerly & Stauder (1958), Stauder & Byerly (1958), and Stauder (1959). Honda (1957), who used *S* waves already many years ago, has published a comprehensive survey of the earthquake mechanism problem, including an extensive bibliography.

There are three main problems involved in such investigations:

(1) The relation between focal mechanism and vibration of *S* waves, as they leave the source. This problem has been dealt with by Monakhov (1950), Dehlinger (1952), Gutenberg (1955), Honda (1957), Keylis-Brook (1957), Nuttli (1958), Stauder & Byerly (1958), and others.

(2) The changes in the vibration properties during the propagation from source to station. This problem is studied in the present paper.

(3) The relation between vibration of an incident *S* wave and seismograph records. This problem has been treated by Gutenberg (1952), Ingram (1953) and others.

Returning to point (2) above, we can summarize our results as follows. Observations of direct *S* waves are most trustworthy, as there is practically no change in the vibration angle during the propagation through the mantle (a continuously varying medium). Changes are expected to arise as the waves pass the base of the crust towards the station, but as the Moho is only one or two wavelengths from the station, it is likely that the effect is much smaller than it would be at a greater distance from a discontinuity surface. Moreover, the change of the vibration angle on transmission through Moho is quite small, as is evident from Figure 4. On the other hand, in using other transverse waves (*ScS*, *sS*, *SS*, ...) we must necessarily take the changes of the vibration angle on reflexions at the core boundary or at the Earth's surface into account.

Different methods have been used in observations of transverse waves for deducing focal mechanisms, namely:

(1) Observation of the direction of first motion (Heinrich & Haill 1952, Gutenberg 1955, Honda 1957, Byerly & Stauder 1958, and others).

(2) Observation of the direction of vibration by combination of simultaneous amplitudes on two or preferably three matched components (Neumann 1930, Dehlinger 1952, Ritsema 1957, Stauder 1959). This is equivalent to the method using the amplitude ratio *SH*/*SV* (Keylis-Borok 1957).

(3) Observation of amplitude ratios between *P* and *S* waves (Keylis-Borok 1957, Honda 1957).

First-motion observations of *S* waves are generally less reliable than observations of polarization, owing to the already existing motion in a seismogram. Moreover, it must be observed that the direction of initial motion may deviate considerably from the direction defined by the vibration angle (δ), which is related to the fully developed motion after the initial stage.

Seismological Laboratory,
University of Uppsala,
Uppsala, Sweden:

1960 *September.*

References

Adams, W. M., 1958. A study of earthquake mechanism using *S* wave data. *Bull. Seismol. Soc. Amer.*, **48** (3), 201–219.

Byerly, P. & Stauder, W., S.J., 1958. The mechanism at the focus of an earthquake. *Earthquake Notes*, **29** (3), 17–23.

Dehlinger, P., 1952. Shear-wave vibrational directions and related fault movements in Southern California earthquakes. *Bull. Seismol. Soc. Amer.*, **42** (2), 155–173.

Galitzin, B., 1911. Ueber die Schwingungsrichtung eines Bodenteilchens in den transversalen Wellen der zweiten Vorphase eines Bebens. *Bull. Acad. Imp. Sci., St.-Pétersbourg*, 1019–1028.

Geiger, L. & Gutenberg, B., 1912. Über Erdbebenwellen VI. *Klg. Ges. d. Wiss. Göttingen, Nachrichten, Math.-phys. Kl.*, 623–675.

Gutenberg, B., 1952. *SV* and *SH*. *Trans. Amer. Geophys. Union*, **33** (4), 573–584.

Gutenberg, B., 1955. The first motion in longitudinal and transverse waves of the main shock and the direction of slip (Kern County earthquakes 1952). *Division of Mines, San Francisco, Bull.*, **171**, 165–170.

Heinrich, R. R. & Haill, H. K., 1952. On the recorded motion of the *S* phase at Florissant. *Bull. Seismol. Soc. Amer.*, **42** (2), 145–154.

Honda, H., 1957. The mechanism of the earthquakes. *Sci. Rep. Tôhoku Univ., Geophysics*, **9** (*Suppl.*), 1–46.

Ingram, R. E., S.J., 1953. Vibration angle of *S* wave. *Bull. Seismol. Soc. Amer.*, **43** (2), 145–151.

Jeffreys, H., 1926. The reflexion and refraction of elastic waves. *Mon. Not. R. Astron. Soc. Geophys. Suppl.*, **1** (7), 321–334.

Jeffreys, H. & Bullen, K. E., 1940. *Seismological Tables*. (Brit. Ass. Advancement of Science: London) 48 pp.

Keylis-Borok, V. I., 1957. The determination of earthquake mechanism, using both longitudinal and transverse waves. *Ann. di Geofisica*, **10** (1–2), 105–128.

Monakhov, F. I., 1950. On the polarization of transverse seismic waves. *Izv. Akad. nauk SSSR, Ser. Geograf. & Geofiz.*, **14** (6), 501–513.

Neumann, F., 1930. An analysis of the *S* wave. *Bull. Seismol. Soc. Amer.*, **20** (1), 15–32.

Nuttli, O. W., 1958. A method, using *S* wave data, of determining the direction of horizontal forces which produce an earthquake. *Earthquake Notes*, **29** (2), 12–14.

Ritsema, A. R., 1957. On the use of transverse waves in earthquake mechanism studies. *Lembaga Meteorologi dan Geofisik, Djakarta, Verhandl.*, **52**, 1–32.

Stauder, W., S.J., 1959. A mechanism study: the earthquake of October 24, 1927. *Geofisica pura e app.*, **44** (3), 135–143.

Stauder, W., S.J. & Byerly, P., 1958. Nodal lines for *S* waves. *Earthquake Notes*, **29** (3), 24–30.

S and the Structure of the Upper Mantle

I. Lehmann

"I have been insisting for about 20 years that the claim of finality for any scientific inference is absurd."—*H. Jeffreys*, The Times of *P*, *S* and *SKS*, and the velocities of *P* and *S*, *Mon. Not. R. Astr. Soc. Geophys. Suppl.*, **4**, 1939, 506.

Summary

The European as well as the north-eastern American observations of *S* at small epicentral distances indicate the presence of a low velocity layer. In Europe its upper boundary seems to be at a depth of about 140 km. Since late *S* phases are observed at epicentral distances down to about 10° there is likely to be an abrupt increase of velocity (as well as of velocity gradient) at the lower boundary of the layer at about 220 km depth. Late *S* phases beyond 20° can be accounted for if a further strong increase of velocity gradient at a greater depth is assumed.

1. Introduction

It is well known that in shallow and not very deep earthquakes, *S* at epicentral distances smaller than about 25° causes difficulty. The "true" *S* is mostly weak or completely missing and a later, larger phase is often read for it. This seems to be true for most regions, and, using I.S.S. data, Lee (1938) demonstrated that at small epicentral distances and especially from about 5° to 14°, *S* travel times failed to concentrate on a mean value. Gutenberg, who always considered amplitude variation in relation to the shape of the time-curve, saw that both *P* and *S* were relatively weak at small epicentral distances, and in several publications he drew attention to the fact and suggested velocity functions that would account for it (Gutenberg 1926, 1959; for further references consult the latter publication). Wadati & others (1933) have remarked on the ambiguity of *S* and mentioned the presence of a large phase, seemingly not the true *S*. When the time-curve for *S* of a surface shock was constructed the curve from about 16° onwards was first obtained by reduction of the travel times of a deep focus shock and it was afterwards extended to small epicentral distances on the assumption that *P* and *S* travel times were always in the same ratio. It was a similar though more refined procedure that Jeffreys (1939) used when it had been found that the time-curve could not be constructed directly from observations.

Explosion results showed that the Jeffreys–Bullen (J.B.) transmission times for longitudinal waves were in error at small epicentral distances and European travel times have been revised in recent years (Jeffreys 1952, 1954, 1958). The *P* curve was found to be very nearly straight up to 15°. From this distance onwards the time-curve had a considerable bend. The upper part of the time-curve is found to be explicable as that of a wave refracted in a surface of discontinuity placed at a depth of 220 km or thereabouts (Lehmann 1959).

When a revision of the *S* transmission times for short epicentral distances is to be undertaken we have to consider that the velocity of *S* like that of *P* has originally been taken too small just below the Mohorovičić discontinuity. From the J.B. tables the initial velocity of the direct wave *Sd* was found to be 4·35 km/s while recent explosion work has yielded velocities of about 4·7 km/s. The results have no great certainty and the velocity may not be quite so high, but there is no doubt about its being greater than the J.B. value.

2. N. American observations

In north-eastern America *Sd* was found to be better recorded than in Europe and perhaps better than in most parts of the world (Lehmann 1955). In the few larger earthquakes that occurred in north-eastern America *Sd* was a clear and well-recorded phase and the time-curve could be represented by the straight line

$$t = a\,\text{s} + \Delta^\circ \times 24{\cdot}07\,\text{s/deg}$$

with *a* varying a little from one earthquake to another. At 14° epicentral distance the time-curve broke off; although *S* was clearly recorded at about this distance it did not appear at greater distances. From about 20° onwards the *S* curve was found to be a J.B. curve at its original height. At intermediate distances there were only few observations, but the travel times fitted on approximately to the J.B. curve that at 14° is 13 s above the *Sd* line as determined.

The break in the time-curve and the delay of the second branch against the first one indicates the presence of a low velocity layer. At the depth grazed by the ray emerging at 14° there is a sudden or a strong gradual decrease of velocity with depth. When this was realized it seemed at first as if the structure in north-eastern America of the upper mantle must differ radically from the structure in other regions having a low velocity layer at a different depth. In the course of further studies it was found that this was not the case.

We may ask at what depth the low velocity layer of north-eastern America is likely to be. The observations at hand are not very numerous and it is not possible to construct a time-curve with the accuracy that would be needed if the velocity distribution were to be derived from it. The *Sd* curve is approximately a straight line, but it probably has a slight curvature corresponding to some increase of velocity with depth since the phases are so well recorded. It was assumed tentatively that the low velocity layer began at 120 km depth. Then the ray emerging at 14° had to have its deepest point at this depth and this would happen if the velocity increased downwards from the Mohorovičić discontinuity by about 0·05 km/s, from, say, 4·60 to 4·65 km/s. The time-curve calculated on this assumption has a slight bend that would not be shown by earthquake observations. Whether this velocity increase would be strong enough to produce clear *S* phases in earthquakes not excessively strong is a question it may be difficult to answer. If we have to assume the velocity gradient to be higher, the ray emerging at 14° will be more strongly bent and go deeper down so that we shall have to take the low velocity layer at a greater depth. It was found subsequently that in Europe the low velocity layer had to be taken deeper down, at a depth of 140–150 km. If in north-eastern America it is at the same depth, the velocity increase above the layer will be greater.

A low velocity layer for *S* is also indicated by surface wave observations as found at the Lamont Geological Observatory. During my stay there in the winter

of 1958 J. Dorman and M. Landisman were calculating dispersion curves for Rayleigh and Love waves on various velocity assumptions trying to obtain agreement with those empirically determined. It was of course important that the velocity assumptions they made should yield also acceptable time-curves and therefore co-operation began, M. Landisman undertaking to calculate time-curves on velocity assumptions I suggested, on the Columbia University automatic calculator, IBM 650. This was an invaluable help to me for, although the computations are quite simple when the velocity law $v = v_0(r/r_0)^{-k}$ is adopted, yet they are highly time-consuming.

The velocity was taken to decrease abruptly at the upper boundary of the low velocity layer, at first assumed to be at 120 km depth. There is, as far as I can see, nothing to show whether the decrease is abrupt or gradual, so for simplicity the velocity was taken to be constant in the low velocity layer and to increase again abruptly at the lower boundary at 220 km depth where a discontinuous increase of the *P* velocity also has been taken to occur.

The ray for which $dt/d\Delta = r_1/v_1$ for the upper boundary grazes this boundary. With $dt/d\Delta$ decreasing the rays enter the layer and they are totally reflected at the lower boundary until $dt/d\Delta$ assumes the value $r_1/v_1 \sin i = r_2/v_2$ of the lower layer. For smaller $dt/d\Delta$ they pass into the layer. The "first" totally reflected ray emerges at a greater epicentral distance than the "last" direct ray as appears from Figure 1.

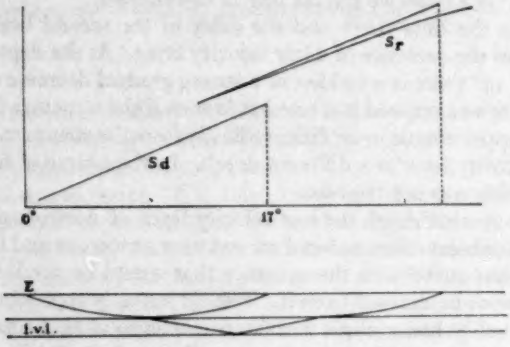


FIG. 1.—Ray paths and time-curves in presence of low velocity layer.

The time-curve of the reflected wave is retrograde and it is some way above the time-curve of the direct wave, the distance between the two lines depending on the thickness of the low velocity layer and the velocity in it. The smallest distance at which we have a totally reflected wave depends also on the velocity below the discontinuity. The refracted wave, *Sr*, begins to emerge at this distance, but it carries very little energy at first. With increasing velocity in the lower layer the intensity of the ray increases and so the *Sr* phase increases with distance.

A velocity distribution of the kind here described was considered by Dorman & others (1960) and depicted in their Figure 2. The crust was taken to be 35 km thick and the velocity in it 3·55 km/s. In the uppermost mantle, down to 120 km depth, the velocity increases from 4·60 to 4·65 km/s. In the low velocity layer the velocity is 4·30 km/s and at the lower boundary it increases to 4·70 km/s. It increases uniformly down to 410 km depth where there is an abrupt

increase from 4·95 to 5·15 km/s. It again increases uniformly down to 600 km where it is 5·65 km/s or practically the same as the J.B. value. On the whole there is no great departure from the J.B. velocities below 220 km depth.

The *Sd* line calculated on the above assumptions has the slope 23·8 s/deg. Originally the north-eastern American *Sd* travel times were fitted to a line of slope 24·0 s/deg, but they have no great precision and they fit on to the new line as well or better. The calculated *Sr* travel times are 16 s delayed against *Sd* at 14°. Around 20° the time-curve has the same shape as the J.B. curve but it is slightly below.

Rayleigh wave dispersion curves calculated on the above assumptions were found to be in good agreement with observations (*loc. cit.*). Love wave dispersion curves were calculated on similar assumptions by Landisman (unpublished) and by Jobert (1960), but here the agreement with observations was not quite so good.

3. European observations

It is north-eastern American observations of *S* that have been considered so far. In Europe we have many more observations of *S* at short epicentral distances, but the so-called true *S* is usually weak even in the stronger shocks and it is difficult to obtain good observations of it. Most European shocks are not large. Occasionally large shocks occur in Italy and also in Greece, but the Greek surface

FIG. 2.—Paris records of Azores earthquake $\Delta = 17°{\cdot}5$.

shocks do not seem to give useful records of *P* and *S*. Galanopoulus (1953) has remarked on the fact that both *P* and *S* are marked by groups of oscillations increasing from small and indefinite beginnings. I have made a study of the European records of one such shock and I found it difficult to fix on corresponding onsets in the records of the different observatories.

In spite of the fact that improved data have accumulated since the time when the J.B. travel times were determined it may still be impossible to construct an *Sd* time-curve with any accuracy from the observations of shallow shocks. The *S* travel times of individual shocks usually scatter a great deal, and since the time-curve as we have it is likely to be in error, we cannot pick out the *Sd* phases by comparison with it. It is only when the earliest readings of a number of stations within a certain distance range are sufficiently consistent to indicate a time-curve clearly that the phases with some confidence can be identified as *Sd*, and this does not seem to happen very often.

Searching the I.S.S. I found one earthquake in which unusually many of the *S* travel times reported seemed to be those of *Sd*, and this is the earthquake of 1950 September 5 in the Gran Sasso in Central Italy. There are 37 reporting stations at distances between 2° and 10° epicentral distance. The *P* and *S* travel times of 27 of these together with some for greater distances are tabulated in Table 1. Some of the stations omitted have either abnormally large negative *S* residuals or no *S*; others have erroneous *P* readings and this casts a doubt on the reliability of their *S*. There are no useful *S* readings at distances greater than that of Moscow (20°·4). Some of the stations have two *S* readings; those in brackets are in the I.S.S. tabulated as additional readings. The time of occurrence has been taken 2 s later than for the I.S.S.

The residuals of *P* are with reference to the tables I have calculated (Lehmann 1959, p. 389); from 22° onwards the travel times are the same as the J.B. times with 3 s subtracted from them. For *S* the $O–C_1$ are the residuals against the J.B. $S-3$ s. In the first column we have placed the residuals of all the "early" readings, regardless of whether in the I.S.S. they have been tabulated as *S* or as additional readings. At the smallest distances these residuals scatter a great deal more than at greater distances, possibly because surface waves interfere and make it difficult to pick the *S* phase. From 6° to 10° the residuals are consistently negative. The first of the $O-C_2$ columns contains the residuals of the same travel times against

$$t = 14\text{s} + \Delta^\circ \times 23{\cdot}8\,\text{s/deg}.$$

They are small and have no systematic trend and the line will be taken to represent the *Sd* time-curve. A line of the same slope was found to fit also the *Sd* of north-eastern America. It corresponds to a velocity of about 4·63 km/s in the uppermost mantle, and this velocity is considerably greater than that found from the J.B. tables.

Beyond 10° we have no travel times fitting on to our *Sd* line so there is a shadow for the *Sd* waves of this earthquake beginning at 10°. But we cannot conclude that in other earthquakes there will also be a shadow beginning at this distance. The Gran Sasso earthquake is not large, and the *Sd* waves fade out with distance. We shall see presently that in a very large earthquake *Sd* waves were recorded at greater epicentral distances. The Copenhagen record from an epicentral distance of 13°·1 is the only one I have seen and it is rather small. There is some movement at about the time when *S* should appear, but there is no clear *S* phase; the Copenhagen bulletin has no *S*. When for north-eastern America the *Sd* curve was taken to break off at 14° it was because *Sd* was a clear phase up to that distance and then appeared no more.

It seems necessary to mention here that Di Filippo & Marcelli (1951) made a study of the Gran Sasso earthquake and that they found the *Sd* travel times for distances up to 13° to fit on to a straight line of slope 25·31 s/deg or a slope much

greater than the one here found and greater even than the mean slope of the J.B. *S*-curve. They determine a slightly different epicentre but that does not alter the travel times greatly. In Table 1 the negative residuals against the J.B. times indicating a smaller slope are at distances from 6° to 10°; 13 travel times have large negative residuals. In the study mentioned 8 of these were not considered and for 4 of the others the authors, reading the records themselves, found different travel times. They continued their line to 13° including a reading at De Bilt and also one at Copenhagen where I fail to see an *S*.

We shall maintain here that an *Sd* line of slope of about 23·8 s/deg has been determined and in other shocks we shall find *Sd* travel times fitting on to the line. It seems, however, to be a rare occurrence that an *Sd* line is clearly indicated in any one shock.

In Table 1 most of the *S* travel times that are not *Sd* times have positive residuals against the J.B. *S*–3 s. They scatter greatly and evidently are not all travel times of one and the same wave. From 13° onwards, however, the scatter is small and a time-curve is indicated that at 20° merges into the J.B. *S*–3 s. Beyond 20° *S* of this shock is poorly recorded, but we expect the European *S*-curve to continue as the J.B. *S*–3 s so that, with *P* being the J.B. *P*–3 s, the *S*–*P* interval is preserved. For north-eastern America the *S* curve was some seconds higher from around 20° onwards and the *S*–*P* exceeded the J.B. *S*–*P*.

But here as for north-eastern America the *Sd* curve is a separate branch not connecting with the rest of the *S*-curve. If prolonged beyond 10° it intersects the J.B. *S*—3 s curve at 21°. At smaller distances the other *S* phases are greatly delayed against *Sd* as shown by the residuals in the second O–C_2 column of Table 1. A structure accounting for north-eastern American travel times has been indicated and it appears that a similar structure will account for *Sd* and also for *S* as observed from about 13° onwards, the waves associated with the upper branch of the time-curve being explainable as waves refracted in a discontinuity surface at depth about 220 km, and the delay being brought about by a low velocity layer above.

Some of the observations at distances smaller than 13° are possibly of the refracted wave *Sr*, but all the observations are not of this phase. It has been mentioned that surface waves possibly interfere with *Sd* at the smallest distances, and at somewhat greater distances we can expect higher mode surface waves to arrive later, but not very much later, than *Sd*; with their relatively short period their appearance may be similar to that of *S* (see Lehmann & Ewing 1960). The O–C_3 of Table 1 are residuals against $t = \Delta° \times 27{\cdot}8$ s/deg corresponding to surface velocity of 4 km/s which is the velocity of second mode surface waves of period 11 s–12 s. Many of the travel times fit on very well to this line, but we should have to examine the records to see whether the onsets read mark beginnings of surface wave trains of the appropriate period.

The readings of some of the observatories bear witness of the difficulty there sometimes is in fixing on an *S*-phase. Stuttgart (not in the table) at an epicentral distance of 6°·9 reads *S* at 2 min 30 s (residual −32 s), and there are additional readings at 2 min 46 s, 54 s, 3 min 00 s, 06 s, 18 s. The time of *Sd* should be 2 min 58 s. Warsaw at 10°·9 also has several readings.

Three observatories, Jena, Collmberg and Bucharest, have read phases not much later than *Sd*, two of them in addition to *Sd*. They probably mark an increase of the *Sd* oscillations. On 1951 August 8 there was another shock from the same epicentre. In this there were several such slightly late *Sd* readings with no

Table 1

Earthquake of 1950 *September* 5, 4 *h* 9 *m* 2 *s*. *Epicentre* 42°·6 N, 13°·5 E

Station	Δ deg	*P* min	*P* s	O–C s	S min	S s	O – C_1 s	O – C_1 s	O – C_2 s	O – C_2 s	O–C_3 s
Firenze	2·0	0	35	0	0	57	−2				
Prato	2·2		34	−3	0	58	−6				
Bologna	2·5		44	2	1	17	6				
Trieste	3·1		47	−3		26	0				
Salo	3·7	0	56	−2		40	−2				
Pavia	4·1	1	04	1	1	50	−2				
Messina	4·7		11	0	2	12	5				
Belgrade	5·5		26	4		42		15		17	9
Zürich	5·9		27	−1		37	0		3		
Wien	6·0		33	4		40 [66]	0	26	3	29	
Budapest	6·3		28	−5		56 [43]	−4	9	−1	12	1
Ebingen	6·4		(41)			47	−3		1		
Neuchâtel	6·4		34	0		43	−7		−3		
Basel	6·5		36	0	2	49	−3		0		
Strasbourg	7·2		46	1	3	04	−6		−1		
Sofia	7·2		49	4		42		32		37	
Cheb	7·5		48	−1		29 [14]	−3	12	1	16	0
Prag	7·5		53	4		12 [32]	−5	15	−1	19	3
Clermont-Ferrand	8·1		56	−2		27 [42]	−5	10	0	15	− 3
Skalnate Pleso	8·1	1	59	1		29	−3		2		
Jena	8·4	2	06	4		40 [34]	−6		0	6	
Collmberg	8·7	1	59	−7		46 [42]	−5	−0	1	5	
Athens	9·1		—		3	50	−7	1	−1		
Bucharest	9·3	2	18	4	4	02		0		7	
Potsdam	9·8		20	−1		34		20		27	2
Paris	9·9		23	1		10 [27]	−7	10	0	17	− 8
Tortosa	9·9		21	−1		21		4		11	
Warsaw	10·9		36	0		48 [54]		7		15, 21	
De Bilt	11·0		—		4	46		2		10	
Alicante	11·5		52	8	5	16		20		28	− 4
Istanbul	11·7		—			15		14		23	−10
Kishimev	11·7	2	52	6		—					
Kew	12·9	3	04	1		53		23		32	− 6
Copenhagen	13·1		10	5		—					
Toledo	13·5		11	0	5	48		4		13	
Malaga	15·0		34	3	6	26		6		15	
Yalta	15·1	3	33	1	6	27		5		14	
Uppsala	17·5	4	04	0	7	23		5		12	
Aberdeen	17·6		06	1		29		9		16	
Lisbon	17·6		14	9		24		4		11	
Helsinki	18·9		22	2		52		2		8	
Helwan	19·1		22	−1	7	56		2		7	
Pulkovo	20·1		34	0	8	13		−3		1	
Moscow	20·4	4	38	1	8	21		−1		1	

Note.—For explanation of residuals, see text.

preceding *Sd* to fit our line, and they prevented the *Sd* line as we have determined it from being clearly apparent.

An earthquake has been found in which *Sd* phases are clearly recorded out to an epicentral distance of nearly 19° and this is the large Azores earthquake of 1931 May 20, magnitude 7·1 according to Gutenberg & Richter (1949). Its epicentre is taken at 37°·6 N 16°·0 W between the Portuguese coast and the Azores. In a study of this earthquake early *S* phases were read that could not be explained at the time (Lehmann 1934, p. 33). The travel times are now found to fit in to our *Sd* line of slope 23·8 s/deg. The readings as they were given in the earlier publication and their residuals against our line are tabulated in Table 2. The *Sd* phases are

Table 2

Earthquake of 1931 May 20, 2 *h* 22 *min* 53 *s*. *Epicentre* 37°·6 N, 16°·0 W

Station	Δ	*Sd*		O–C
	deg	min	s	s
Granada	9·9	4	06	−3
Alicante	12·3	5	10	4
Tortosa	13·2	5	30	2
Barcelona	14·5	5	58	−1
Paris	17·5	7	09	−1
Kew	17·8	7	17	0
Liverpool	18·2	7	26	−1
Stonyhurst	18·7	7	37	−2

small but quite clear as seen in the Paris record of Figure 2 from an epicentral distance of 17°·5. With constant velocity in the uppermost mantle the ray to 19° would have its deepest point at about 115 km depth; with slightly decreasing velocity it would not go quite so deep. The structure in the epicentral region is not known but it is perhaps likely to be intermediate between continental and oceanic and to have the low velocity layer at a smaller depth than under the continents. In addition to *Sd* phases late *S* phases have also been recorded in the Azores shock.

The earthquake of 1933 April 23 in the Dodecanese had only few observations at short epicentral distances but *Sd* was read at 4 stations at distances smaller than 10°. All other *S* phases at distances up to 20° were late, and late *S* phases were also recorded beyond 20° as found in a previous study of this earthquake (Lehmann 1953).

The Italian earthquake of 1933 July 23 has also been considered. It is a very much larger earthquake than the 1950 one, but there are in the I.S.S. only 8 readings of *Sd* that fit on to our line. The greatest epicentral distance at which *Sd* has been reported is 13°·0 (Hamburg). Some of the stations have no *S* but most of them have late readings. The residuals of these scatter far more than in the other earthquakes considered. This and the lack of *Sd* readings may partly be due to the fact that instrumental equipment was poorer in 1930 than it was in 1950, but the *S* phase may also be less well developed and not so clear. At Copenhagen at a distance of 14°·7 where the late *S* usually has begun to increase the phase is quite small. Much larger second mode surface waves arrive about 20 s later on the E component record and somewhat later on N (see Lehmann & Ewing 1960, Figure 6).

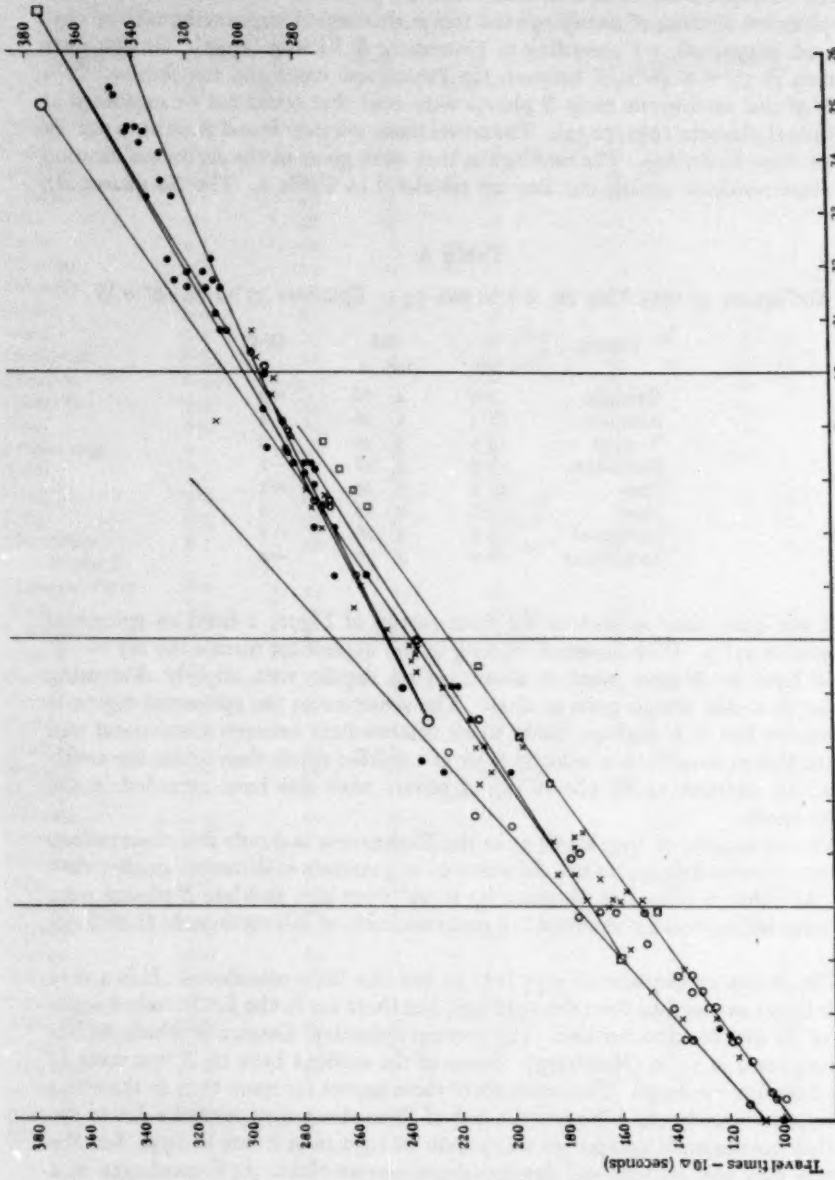


FIG. 3.—4 shallow earthquakes. Observed and calculated travel times of *S*.

The time-curves of Figure 3 were calculated on velocity assumptions similar to those on p. 126. The *Sd* velocity was taken a little smaller, and the slope of the *Sd* line is 24·0 s/deg. In the low velocity layer the velocity is as before 4·30 km/s, but at the 220 km discontinuity it increases to 4·75 km/s and at 410 km the velocity increase is from 4·90 to 5·05 km/s. At 600 km the velocity is as before 5·65 km/s. Above the *Sd* line are the time-curves of the waves reflected on and refracted in the two discontinuity surfaces. At distances up to nearly 20° the lowest of the curves is the Sr_1 time-curve of the wave refracted in the 220 km discontinuity and it continues beyond 20° with no great curvature. The Sr_2 time-curve of the wave refracted in the 410 km discontinuity crosses it between 19° and 20°; beyond 20° it is identical with the J.B. *S*—3 s curve. To the left there is the line of slope 27·8 s/deg corresponding to surface velocity 4 km/s. The times of the figure are travel times minus 10Δ s.

The points plotted in the figure are from our four earthquakes. Open circles, squares, dots and crosses are from the 1950, 1931, 1933 and 1930 earthquakes respectively. The points of the late Azores phases have not been plotted, but with few exceptions they would have been very close to points in the figure.

There are many points on or near the *Sd* line, most of them from the 1950 earthquake. The squares of the Azores earthquake are a little below the line and, as we have seen, they fit on to a line of slope 23·8 s/deg. It seemed at first as if the points of the other earthquakes fitted better to a line of slope 24·0 s/deg, but it was found later not to be so, and the slope 23·8 s/deg was adopted.

Several points are close to the line of slope 27·8 s/deg, and these may be due to second mode surface waves.

For distances smaller than 20° there are many points on or close to the Sr_1 curve and there are also some beyond. Most of the corresponding phases have been reported as *S*. The points follow the curve closely enough to make it seem safe to conclude that they are explainable on the kind of assumption made; they can be taken to be due to waves refracted in a layer of high velocity at about 220 km depth and to be delayed by a low velocity layer above. There are points on the Sr_1 curve from about 10° onwards. The appearance of refracted (or reflected) waves at so short an epicentral distance indicates that the velocity increase at 220 km depth is discontinuous. A gradual increase could also produce a retrograde branch of the time-curve, but this could not go so far back. There is a considerable scatter of points between 10° and 12° but it is due chiefly to the 1930 earthquake the *S* phases of which do not seem to be very clear.

A second, deeper discontinuity in the velocity was assumed because late *S* phases often are observed beyond 20°. This was noticed by Jeffreys & Bullen (1935, p. 41). It has been pointed out in some of my earlier studies and it was found that straight lines of slope of about 21 s/deg could be fitted to the travel time points of the late phases for distances both smaller and greater than 20° (Lehmann 1934, Figures 3 and 4; 1953, Figure 4). However, no structure was found that would account for these lines. Now, taking the velocity gradient below the 220 km discontinuity to be small enough for the Sr_1 curve to continue beyond 20° with no great curvature, we obtain an explanation of the late phases at those distances. But then we cannot account for the phases associated with the strongly bending time-curve merging into the J.B. *S*–3 s without assuming that at some depth there is a strong increase of velocity gradient and, perhaps, of the velocity itself.

In a recent report on the results obtained for *P* and *S* velocities in the upper mantle (Lehmann 1960) a deeper discontinuity was not mentioned. This was because, on re-examining some records, I did not feel convinced that two distinct phases were actually present and we see also that the points plotted in Figure 2 scatter greatly. However, late phases are often read, but no time-curve has been determined and further investigation is needed.

This also applies to distances smaller than 20°. Here most of the *S* phases reported are late phases. For these no time-curve has been determined and it would probably be found impossible to determine time-curves for them with any accuracy using I.S.S. data. Those we have considered may be said to confirm in a general way that the time-curves of the phases are of the type here proposed, but they do not determine the time-curves in detail. In any one earthquake the residuals scatter a great deal and they also differ from one earthquake to another. This may partly be due to the fact that no special attention has been directed towards these phases. The routine observer will be looking for an *S* phase and some will fix on the first trace of a phase that could be *S* and neglect the subsequent movement while others will expect *S* to be a more pronounced phase and prefer to read the later, larger phase for *S*. No care is probably taken to distinguish between the onset of a body wave and that of a train of short-period surface waves, and the distinction is, in fact, by no means always clear. Special studies of the *Sr* phases are highly needed.

In Japan special studies have been made of *P* and *S* phases as recorded at epicentral distances from about 10° to a little beyond 25° (Nishimura & others 1958). Travel times, amplitudes and periods were measured in the records of 23 observatories all equipped with instruments of the same type. An *Sd* phase distinguishing itself by a period smaller than 2 s was from 13° onwards followed by a much larger Sr_1 phase of greater period. The two time-curves intersected at 18°. The amplitude of Sr_1 decreased with distance and was quite small at about 19° from which distance onwards it was followed by an Sr_2 phase. The time-curves of the two *Sr* phases intersected at about 24°. The amplitude of Sr_2 was also large where the phase first appeared and it decreased with distance.

Although we do not possess such precise results for Europe we can say definitely that the picture here given does not cover European findings. Our Sr_1 phases are not large at the smallest distances and they increase with distance. As to Sr_2 our information is vague, but very large late *S* phases do not seem to appear around 20° whereas late phases sometimes are very large at somewhat greater distances (see Lehmann 1953, p. 480, Figure 3). Our *Sd* line intersects the *Sr* curves at a distance greater than 20°.

For Europe results of far greater precision than those we possess are required if we are to determine the velocity distribution in the upper mantle with any precision. Travel times should be determined but we should know also how the periods and amplitudes of the different waves vary. Such knowledge, however, is not easily obtained for Europe where instrumental equipment is not very uniform and most horizontal instruments have no particular good response to *S* waves at short epicentral distances. We have the short-period vertical seismographs for *P*, but no special effort has been made to obtain good records of *S* at small distances. The 200 kg Wiechert horizontal seismograph used in Japan would not have sufficient magnification for European shocks, but the *Lg* seismograph having its maximum magnification for waves of period 3 s–6 s would undoubtedly produce far better observations of *S* than those now obtainable.

In the Copenhagen records of the earthquake of 1936 October 18 in the Italian Alps, we have an example of the influence of seismograph characteristics on the *S* records obtained. *Sd* is barely visible in the records of the long-period instruments (see Lehmann & Ewing 1960, p. 12 and Figure 4), whereas it is a distinct phase in the record of a Wood–Anderson instrument operating at the time. Its constants were not accurately known, but it is apparent in the record that it responded very well to waves of short period. A clear and relatively strong phase appears 21 s later, also in the records of the other instruments, and this is the phase that was reported as *S*. It seems to be the *Sr* phase here appearing at so short an epicentral distance.

4. Deep focus earthquakes

So far I have considered only shallow shocks. On my velocity assumptions the travel times of deep shocks having their foci below the 220 km discontinuity will be practically the same as those of the J.B. tables, for below the discontinuity the velocity distribution is only slightly altered and the mean velocity in the upper layers is nearly the same. But the travel times of the earthquakes originating in the upper layers will be more or less affected. It is therefore of interest to see whether or not observed travel times of shocks having their foci in the upper layers are in accordance with the travel times calculated on our velocity assumptions.

In a recent investigation the travel times of the *P* waves of six Rumanian earthquakes having a common focus at 130 km depth were considered (Lehmann 1959). For 31 observatories at distances up to 24° the *P* travel times of the various shocks were found to be in excellent agreement and well-determined means of the travel times of individual observatories were used.

The *S* travel times of the shocks are not in such good agreement but scatter a great deal. *S* observations usually have smaller accuracy than those of *P* because the movement does not start from rest and, indeed, often emerges from a strong background movement. This, however, was not the only reason why the observations scattered. We know that there is often more than one *S* phase observed at distances up to and a little beyond 20°, and it was evident that it was not always the same phase that had been selected and reported as *S*. The I.S.S. travel times of some of the observatories did not all concentrate on one and the same mean value but indicated that in some of the earthquakes one phase had been read while another phase had been read in the others. It was not always easy, however, to separate the phases when only the I.S.S. travel times were considered since the observations of each of them scattered. It was found necessary to examine the records of one of the shocks to see what phases were present.

The records of 24 European stations of the earthquake of 1940 October 22 were used. In several of the records more than one *S* phase was very clearly present. Often the mean of the I.S.S. travel times was in fair agreement with the travel time of one of these phases, but sometimes, when the I.S.S. travel times scattered too much for a mean to be formed, they were close enough to the travel times of different phases in the 1940 records to indicate that these were the phases read. Using the I.S.S. data for all the earthquakes and in addition my own readings of 1940 October 22, I determined the travel times of individual stations. Some of these were quite well determined means while others were less good; they are marked in Figure 4 by filled-out and open circles respectively. Some observations were not obtained in more than one shock; the corresponding travel times are marked by crosses.

The time-curves of Figure 4 were calculated on velocity assumptions differing somewhat from those used for surface shocks. The velocity in the uppermost mantle is 4·63 km/s and the low velocity layer is taken to begin at 150 km depth, but it extends as before to 220 km depth where the velocity increases abruptly. The velocity in the layer is only 4·10 km/s and so the delay of the *Sr* phases is approximately the same as before. We have again a second discontinuity at 410 km depth.

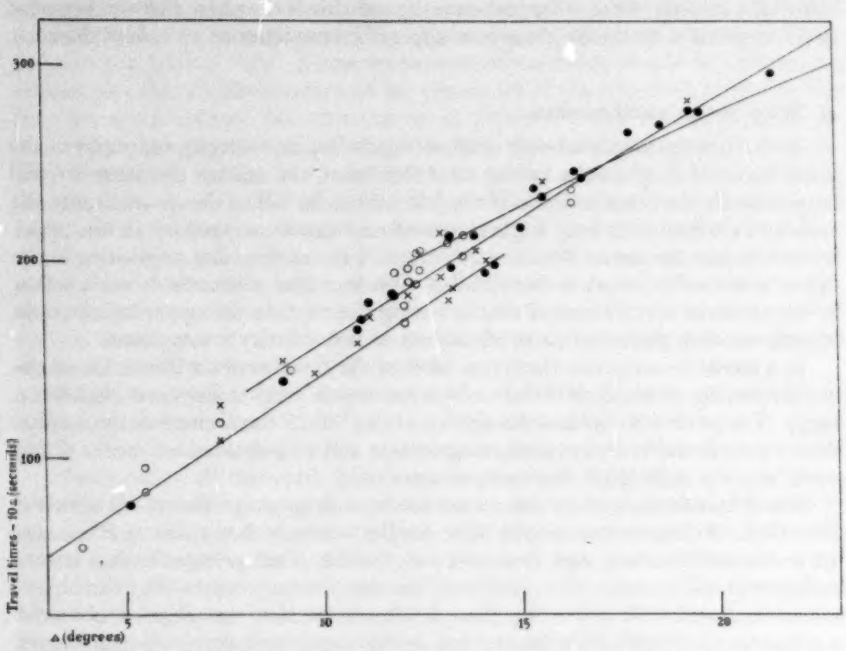


FIG. 4.—6 Rumanian earthquakes, depth 130 km. Observed and calculated travel times of *S*.

The slope of the *Sd* line for a surface focus will be 23·8 s/deg. For 130 km depth the slope of the *Sd* line is 23·4 s/deg at the inflexion point at 10° epicentral distance, and the mean slope is approximately the same from about 5° to 15°. There are a few well-determined points and some more doubtful ones on the *Sd* line of Figure 4. The two points on its upper end are those of Helsinki and Pulkowo at epicentral distances 14°·4 and 14°·2. I have seen the Helsinki records of three of the shocks and have found the phase to be clear and its amplitude surprisingly large. It is obviously the *Sd* phase that has been recorded and this phase could not be recorded at a distance greater than that of the inflection point, 10°, if the low velocity layer began at the depth of focus. That is why the low velocity layer has now been taken to begin at 150 km depth, but it could perhaps have been taken a little higher up.

Most of our points are grouped around the Sr_1 curve above the *Sd* line. The two *Sr* curves intersect at about 17° and also beyond this distance there are points close to the Sr_1 curve seemingly confirming that it continues with no great curvature. There are no points on the Sr_2 curve beyond 17° in the figure, but the

travel time points for greater distances fit on to its continuation. We have some well-determined points on the lower part of the Sr_2 curve; the corresponding phases are very clear in the records examined.

Usually we do not for *P* observe double phases around 20° corresponding to two different *Pr* curves. In the Azores earthquake, however, a second phase following closely upon the first one was read in 7 records at distances from 20°·4 to 21°·3 and the phase was interpreted as *Pd* (Lehmann 1934, p. 19). It now seems as if it may be Pr_1 while the earlier phases belong to a Pr_2 branch that continues beyond.

The travel times of *P* were found to be consistent with a structure that had no low velocity layer, and although it is physically possible for the velocity of *S* to decrease while that of *P* does not, yet it seems more likely that the *P* velocities are somewhat affected in the layer in which the *S* velocities are low. The velocity distribution for *P* was therefore altered so as to include a layer of slightly smaller velocity extending from 150 to 220 km depth, and it was found possible to do this without altering the travel times appreciably.

5. Conclusion

In conclusion it can be said that a low velocity layer in the mantle is strongly indicated by observations of *S* phases at epicentral distances up to about 25°. In Europe the upper boundary of this layer is at a depth greater than the focal depth of the Rumanian earthquakes considered, and this depth is not likely to be smaller than 130 km. While no indication of the character of the upper boundary of the low velocity layer has been obtained, it is found that at the lower boundary, taken to be at 220 km depth, there is likely to be a discontinuous increase of velocity, for late *S* phases with travel times as those calculated on this assumption appear at small epicentral distances and their amplitudes increase with distance. The presence of a deeper discontinuity is indicated. Special studies of the *S* phases are required if progress is to be made on these lines. More adequate instrumentation for the recording of *S* at small epicentral distances is needed.

Acknowledgments

Dr Mark Landisman has co-operated in the research here presented, calculating travel times on the Columbia University IBM 650 calculator. I am greatly indebted to him for this assistance, for it would not have been possible for me to complete the work without it, and I am grateful to him for the kindness with which he always was ready to help.

My thanks are due also to Professor Maurice Ewing who invited me to come to the Lamont Geological Observatory and who made it possible for me to work there for several months.

Lamont Geological Observatory,
(Columbia University),
Palisades,
New York:

1961 *September*

References

Di Filippo, D. & Marcelli, L., 1951. Uno studio sul terremoto del Gran Sasso d'Italia del 5 settembre 1950. *Ann. Geofis.* **4**, 213–239.

Dorman, J., Ewing, M. & Oliver, J., 1960. Study of shear velocity distribution in the upper mantle by mantle Rayleigh waves. *Bull. Seismol. Soc. Amer.*, **44**, 127–147.

Galanopoulus, A., 1953. On the intermediate earthquakes in Greece. *Bull. Seismol. Soc. Amer.*, **43**, 159–178.

Gutenberg, B., 1926. Untersuchungen zur Frage, bis zu welcher Tiefe die Erde kristallin ist. *Zeit. Geophys.* **2**, 24–29.

Gutenberg, B., 1959. The asthenosphere low velocity layer. *Ann. Geofis.*, **12**, 439–460.

Gutenberg, B. & Richter, C. F., 1949. *The Seismicity of the Earth* (Princeton University Press).

Jeffreys, H., 1939.1. Some Japanese deep-focus earthquakes. *Mon. Not. R. Astr. Soc. Geophys. Suppl.*, **4**, 424–460.

Jeffreys, H., 1939.2. The times of *P*, *S* and *SKS*, and the velocities of *P* and *S*. *Mon. Not. R. Astr. Soc. Geophys. Suppl.*, **4**, 498–533.

Jeffreys, H., 1952. The times of *P* up to 30°. *Mon. Not. R. Astr. Soc. Geophys. Suppl.*, **6**, 348–363.

Jeffreys, H., 1954. The times of *P* in Japanese and European earthquakes. *Mon. Not. R. Astr. Soc. Geophys. Suppl.*, **6**, 557–565.

Jeffreys, H., 1958. The times of *P* up to 30° (second paper). *Geophys. J.*, **1**, 154–161.

Jeffreys, H. & Bullen, K. E., 1935. Times of transmission of earthquake waves. *Publ. Bur. Centr. Seismol. Intern. A*, 11.

Jobert, N., 1960. Calcul de la dispersion des ondes de Love de grande période à la surface de la Terre. *C.R. Acad. Sci. (Paris)*, **250**, 3693–3695.

Lee, A. W., 1938. The travel-times of the seismic waves *P* and *S*. *Met. Office Geophys. Mem.*, **9**, no. 76.

Lehmann, I., 1934. Transmission times for seismic waves for epicentral distances around 20°. *Geod. Inst. København, Medd.*, no. 5.

Lehmann, I., 1953. *P* and *S* at distances smaller than 25°. *Trans. Amer. Geophys. Un.*, **34**, 477–483.

Lehmann, I., 1955. The times of *P* and *S* in Northeastern America. *Ann. Geofis.*, **8**, 351–370.

Lehmann, I., 1959. Velocities of longitudinal waves in the upper part of the Earth's mantle. *Ann. Géophys. (Paris)*, **15**, 93–118.

Lehmann, I., 1960. Structure of the upper mantle as derived from the travel times of seismic *P* and *S* waves. *Nature (London)*, **186**, 956.

Lehmann, I. & Ewing, M., 1960. On short-period surface waves as recorded in Copenhagen. *Inst. Géodesique Danemark, Mémoires*, Sér. 3, **34**.

Nishimura, E., Kishimoto, Y. & Kamitsuki, A., 1958. On the nature of the 20° discontinuity in the Earth's mantle. *Tellus*, **10**, 137–144.

Wadati, K. & Oki, S., 1933. On the travel times of the earthquake waves (part IV). *Geophys. Mag.*, **7**, 139–153.

Seismic Waves from the Outer and the Inner Core

Pietro Caloi

"It went through and through my head like an earthquake."—*Lewis Carroll, Through the Looking-Glass*

Summary

New formulae for the determination of travel-times of seismic waves in the interior of the Earth. Applications of said formulae to the calculation of travel-times of *ScS*-waves, recorded with exceptional clearness at Tolmezzo, on occasion of Aegean Isles shallow earthquakes (1957 April–June). The problem of the waves reflected by the inner core: manifest examples of *PKiKP*-waves recorded at Tolmezzo.

1. Introduction

Many students of seismology have had the constant preoccupation of determining, with the utmost precision, the behaviour of seismic wave velocity in the interior of the Earth, with particular reference to the mantle. As is known, the general method of Wiechert–Herglotz answers the purpose in this particular case. From the first rudimentary determinations of Wiechert–Zöppritz we progress through the ever-improving approximations of Gutenberg, Witte, Cornelius G. Dahm, Gutenberg & Richter, Jeffreys, and again Gutenberg & others. My intention here is to find out some formulae which, basing themselves upon the last velocity values obtained by Gutenberg, would allow the determination of the travel-time of the seismic waves in the mantle, for known epicentral distances and corresponding to given impulse angles.

2. Travel-time formulae

Considering r as the radius of a certain point of a seismic trajectory, we know that the travel-time dt corresponding to an element of the trajectory is given by

$$dt = \frac{n^2 r\, dr}{\sqrt{(n^2 r^2 - k^2)}},$$

where $n = 1/v$, v being the velocity at the point under examination and $k = nr \cos e$, where e = angle of emergence at that point.

Suffix 0 refers to the values at the surface of the Mohorovičić discontinuity, suffix 1 at the point corresponding to the assumed focal depth, and suffix 2 at the outer core.

Therefore, for the travel-time of a seismic wave (longitudinal or transverse) arriving at the observation station after reflection at the outer core, we have

$$T = \frac{r_0}{v_0}\left(\int_{\rho_2}^{\rho_1} \frac{\nu^2 \rho\, d\rho}{\sqrt{(\phi^2-\alpha^2)}} + \int_{\rho_2}^{\rho_2} \frac{\nu^2 \rho\, d\rho}{\sqrt{(\phi^2-\alpha^2)}}\right), \tag{1}$$

where

$$\nu = \frac{v_0}{v}, \qquad \rho = \frac{r}{r_0}, \qquad \phi^2 = \nu^2\rho^2, \qquad \alpha = \nu\rho\cos e.$$

By means of evident modifications, (1) is obviously also applicable to mixed paths.

Previously (Caloi 1946) a law of seismic wave (either longitudinal or transverse) velocity distribution in the Earth's mantle had been indicated. This law allowed ρ to be expressed as a function of ϕ^2. Our intention here is to express the dependence of ρ on ϕ^2 in a more rigorous way in order to achieve a better approximation.

Table 1 gives the ρ values for various values of v, obtained from the most recent researches on the speed of propagation of seismic waves, both longitudinal or transverse, in the Earth's mantle.

If we consider

$$\rho = \frac{\phi^2}{a + b\phi^2 + c\phi^4 + d\phi^6} \tag{2}$$

we have to determine the constants a, b, c, d, in order to approximate the functions $\rho = \rho(v)$ given numerically.

We note that (2) can be written as follows:

$$\frac{v_0^2}{v^2}\rho = d\frac{v_0^6}{v^6}\rho^6 + c\frac{v_0^4}{v^4}\rho^4 + b\frac{v_0^2}{v^2}\rho^2 + a;$$

from which, given

$$\delta = dv_0^4; \quad \gamma = cv_0^2; \quad \beta = b; \quad \alpha = \frac{a}{v_0^2}$$

we have:

$$\rho/v^2 = \delta(\rho/v)^6 + \gamma(\rho/v)^4 + \beta(\rho/v)^2 + \alpha. \tag{3}$$

Equation (3) shows that ρ/v^2 is a polynomial in ρ/v. It is easy to obtain, from Table 1, the numerical expression of ρ/v^2 as a function of ρ/v.

The method of least squares allows us to determine the values of the constants α, β, γ, δ, from which we obtain the values of a, b, c, d, considering that v_0 is the value that v assumes in the case of $r = r_0$, that is $\rho = 1$.

For the *longitudinal waves* we find:

$$\alpha = +0{\cdot}00157011\,34; \qquad \beta = +0{\cdot}86730457;$$
$$\gamma = +8{\cdot}0454089; \qquad \delta = -346{\cdot}960625;$$

and for the *transverse waves*:

$$\alpha = +0{\cdot}00600129\,98; \qquad \beta = +0{\cdot}78087719;$$
$$\gamma = +4{\cdot}80768275; \qquad \delta = -51{\cdot}283826.$$

On the basis of the v_0 values for longitudinal and transverse waves, we obtain, as an immediate consequence, the corresponding values of a, b, c, d, for the two types of seismic waves.

Table 1

Variation of longitudinal and transverse velocities with radial distance

ρ	v_l km/s	v_t km/s
1	8·1	4·55
0·995269	8·15	4·6
0·988961	8·0	4·4
0·981076	7·85	4·35
0·973190	8·05	4·4
0·957420	8·5	4·6
0·941650	9·0	4·95
0·925879	9·6	5·3
0·910109	10·1	5·6
0·894338	10·5	5·9
0·878568	10·9	6·15
0·862798	11·3	6·3
0·847027	11·4	6·35
0·815487	11·8	6·5
0·783946	12·05	6·6
0·752405	12·3	6·75
0·720864	12·55	6·85
0·689323	12·8	6·95
0·657783	13·0	7·0
0·626242	13·2	7·1
0·594701	13·45	7·2
0·563160	13·7	7·25
0·547390	13·7	7·2
0·544236	13·65	7·2

If we put $\phi^2 = z$, we have

$$\int \frac{v^2 \rho \, d\rho}{\sqrt{(\phi^2 - \alpha^2)}} = \int \frac{a - cz^2 - 2dz^3}{a + bz + cz^2 + dz^3} \frac{dz}{\sqrt{(z - \alpha^2)}}. \tag{4}$$

By approximate methods we obtain the roots of the equation

$$dz^3 + cz^2 + bz + a = 0.$$

They are, with the data at our disposal, for the *longitudinal waves*:

$$\begin{aligned} z_1 &= -0{\cdot}89278418; \qquad z_2 = -0{\cdot}12261503; \\ z_3 &= +1{\cdot}1675373, \end{aligned} \tag{5}$$

and for the *transverse waves*:

$$\begin{aligned} z_1 &= -1{\cdot}644252; \qquad z_2 = -0{\cdot}16825260; \\ z_3 &= +3{\cdot}7532561. \end{aligned} \tag{6}$$

If in (4) we put $u^2 = z - \alpha^2$, we have

$$\int \frac{a - cz^2 - 2dz^3}{a + bz + cz^2 + dz^3} \frac{dz}{\sqrt{(z - \alpha^2)}}$$

$$= -4\left[\int du + \alpha_1 \int \frac{du}{u^2 + \alpha^2 - z_1} + \alpha_2 \int \frac{du}{u^2 + \alpha^2 - z_2} + \alpha_3 \int \frac{du}{u^2 + \alpha^2 - z_3} \right], \tag{7}$$

where

$$\alpha_1 = \tfrac{1}{2} z_1; \quad \alpha_2 = \tfrac{1}{2} z_2; \quad \alpha_3 = \tfrac{1}{2} z_3.$$

Owing to the fact that $z_1 < 0$, $z_2 < 0$ the two first denominators of the second side of the equation (7) are always positive; the third one can have $\alpha^2 > z_3$, $\alpha^2 = z_3$, $\alpha^2 < z_3$.

Therefore, finally we have

$$\int \frac{\nu^2 \rho\, d\rho}{\sqrt{(\phi^2-\alpha^2)}} = -4\sqrt{(\phi^2-\alpha^2)} - \frac{2z_1}{\sqrt{(\alpha^2-z_1)}} \arctan \frac{\sqrt{(\phi^2-\alpha^2)}}{\sqrt{(\alpha^2-z_1)}} -$$
$$- \frac{2z_2}{\sqrt{(\alpha^2-z_2)}} \arctan \frac{\sqrt{(\phi^2-\alpha^2)}}{\sqrt{(\alpha^2-z_2)}} +$$
$$+ \begin{cases} -\dfrac{2z_3}{\sqrt{(\alpha^2-z_3)}} \arctan \dfrac{\sqrt{(\phi^2-\alpha^2)}}{\sqrt{(\alpha^2-z_3)}} + \text{const} & \text{for } \alpha^2 > z_3 \\ \dfrac{2z_3}{\sqrt{(\phi^2-\alpha^2)}} + \text{const} & \text{for } \alpha^2 = z_3 \\ \dfrac{z_3}{\sqrt{(z_3-\alpha^2)}} \ln \dfrac{\sqrt{(z_3-\alpha^2)} + \sqrt{(\phi^2-\alpha^2)}}{\sqrt{(z_3-\alpha^2)} - \sqrt{(\phi^2-\alpha^2)}} + \text{const} & \text{for } \alpha^2 < z_3. \end{cases} \quad (8)$$

Constants z_1, z_2, z_3 take respectively the values given by equations (5) or (6), corresponding to either longitudinal or transverse waves.

In any case, for practical applications, we have $\alpha^2 < z_3$ ($\alpha^2 < 1{\cdot}1675373$ for longitudinal waves and $\alpha^2 < 3{\cdot}7532561$ for transverse waves).

3. Formulae for epicentral distance

Taking $d\theta$ as an element of angular distance, calculated on the Mohorovičić surface, we have

$$d\theta = \frac{\alpha\, d\rho}{\rho\sqrt{(\phi^2-\alpha^2)}}.$$

Corresponding to (1) we have:

$$\Delta = r_0 \alpha \left(\int_{\rho_3}^{\rho_1} \frac{d\rho}{\rho\sqrt{(\phi^2-\alpha^2)}} + \int_{\rho_3}^{\rho_0} \frac{d\rho}{\rho\sqrt{(\phi^2-\alpha^2)}} \right). \quad (9)$$

Also,

$$\int \frac{d\rho}{\rho\sqrt{(\phi^2-\alpha^2)}} = \frac{2}{\alpha} \arctan \frac{\sqrt{(\phi^2-\alpha^2)}}{\alpha} - \frac{2}{\sqrt{(\alpha^2-z_1)}} \arctan \frac{\sqrt{(\phi^2-\alpha^2)}}{\sqrt{(\alpha^2-z_1)}} -$$
$$- \frac{2}{\sqrt{(\alpha^2-z_2)}} \arctan \frac{\sqrt{(\phi^2-\alpha^2)}}{\sqrt{(\alpha^2-z_2)}} + \quad (10)$$
$$+ \frac{1}{\sqrt{(z_3-\alpha^2)}} \ln \frac{\sqrt{(z_3-\alpha^2)} + \sqrt{(\phi^2-\alpha^2)}}{\sqrt{(z_3-\alpha^2)} - \sqrt{(\phi^2-\alpha^2)}} + \text{const},$$

where z_1, z_2, z_3 have the values given by equation (5) for the longitudinal waves and by equation (6) for the transverse waves.

It is evident that (10) and (8) can be used to determine the epicentral distances, corresponding to known values of α, and to determine the corresponding travel-times.

4. Records of *ScS*

In practice, calculations are made to tabulate the formulae as functions of distance and depth. To show their use I am going to work out only four cases—really exceptional ones—of *ScS* wave recordings, obtained at a short distance, for earthquakes at normal depth (or only slightly deeper). The first three are from the Aegean Islands earthquakes of 1957 April 24, 25 and 26.

ScS waves are normally observed only in the case of earthquakes at a great depth (more than 100 km) and cases where such waves can be detected during earthquakes at a normal depth are rare. Moreover, even when they are identified, there is always uncertainty in the interpretation. During my long experience as an observer, I never had the chance to see such evident *ScS* arrivals as those recorded at Tolmezzo, when the earthquakes in the offing of the Dodecanese Islands (Rhodes) took place in 1957 April.

In Table 2 the times at the focus and the epicentral coordinates are listed.

Table 2

The Aegean Islands earthquakes

Date 1957	Source	Origin time G.M.T. h	m	s	Epicentre deg N	deg E	Magnitude
April 24	U.S.C.G.S.	19	10	05	36	28½	6·9 (Uppsala,
	B.C.I.S.	19	10	16	36·3	29·1	Kiruna)
April 25	U.S.C.G.S.	02	25	36	36½	29	7·2 (Uppsala,
	B.C.I.S.	02	25	44	36·3	29·1	Kiruna)
April 26	U.S.C.G.S.	06	33	32	36½	29	5·8 (Uppsala,
	B.C.I.S.	06	33	43	36·3	29·1	Kiruna)

The U.S.C.G.S. considered the focal depths as normal. At first, the B.C.I.S. considered the depths at a depth of about 100 km; afterwards in the *Bulletin Mensuel* the observations were reported without any depth indication.

Markus Båth, who used the records of these earthquakes for a study on the continental dispersion of seismic surface waves, considered the earthquakes to be at a normal depth, following the data given by the U.S.C.G.S. Markus Båth says: "The greater focal depth suggested by B.C.I.S. was not confirmed by the magnitude determinations for Uppsala and Kiruna for either of the shocks, although in the last shock there was a phase 13 s after *P*, possibly *pP*, which would suggest a depth around 70 km" (Båth 1957).

In my opinion—and on the basis of the records obtained at Tolmezzo—the first two earthquakes had different depths: the first, which was the shallower one, originated inside the Earth's crust, while the second was at the base of the crust about (30 km).

The arrival times of *ScS* waves obtained at Tolmezzo, corresponding to these three earthquakes (together with the epicentral distances) were:

	ScS h	m	s	Δ
1957 April 24	19	25	57	1733 km
1957 April 25	02	41	25	1724 km
1957 April 26	06	49	23	1724 km

Taking for α the values 0, 0·1, 0·15, equation (9) (taking into account equation (10)) allows us to determine the corresponding values of Δ and equations (1) and (8) give us the travel-time values. The calculations have been made for a core depth of 2900km, assuming the focus to be at the Mohorovičić discontinuity (calculated at 30km), and $v_0 = 4{\cdot}35$ km/s. The values obtained for T and Δ are:

	T s	Δ km
$\alpha = 0$	917·1	0
$\alpha = 0{\cdot}1$	934·8	1585·4
$\alpha = 0{\cdot}15$	961·7	2456·8

The results of these calculations are given in Figures 1 and 2.

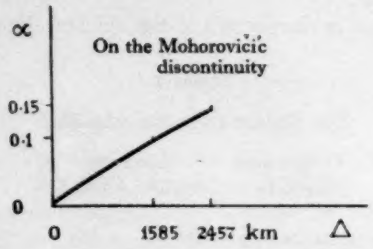


FIG. 1.—$\alpha - \Delta$ for earthquakes at the Mohorovičić discontinuity.

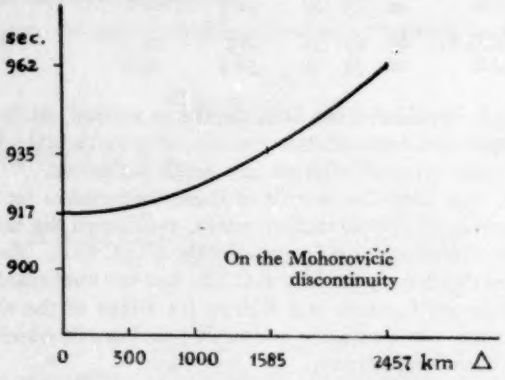


FIG. 2.—Travel-times for earthquakes at the Mohorovičić discontinuity.

Taking into account the corrections due to the passage from the Mohorovičić discontinuity to the outer surface, both for the epicentral distances (+11km) and for the travel-time (+8·5s) the following differences between observed and calculated times are found:

	ScS	
	Observed s	Calculated s
1957 April 24	952	948
1957 April 25	948	947·5
1957 April 26	950	947·5

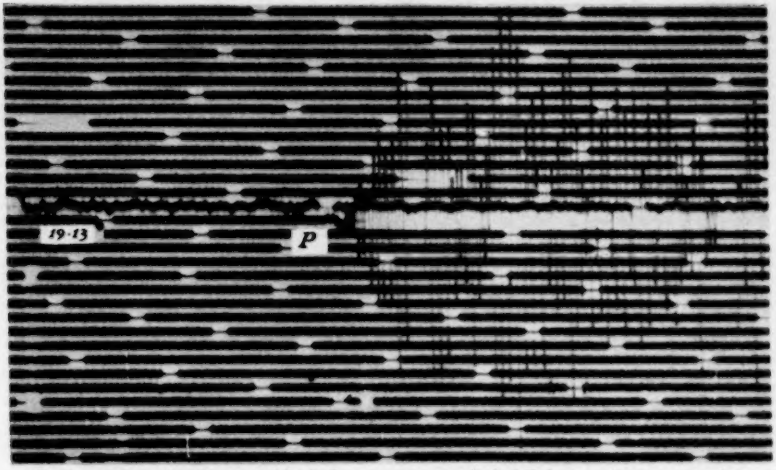


FIG. 3.—1957 April 24 (Aegean Isles). *P*—waves recorded at Tolmezzo (Δ = 1733 km).

FIG. 4.—1957 April 24 (Aegean Isles). *ScS*—and *PKiKP* recorded at Tolmezzo (Δ = 1733 km).

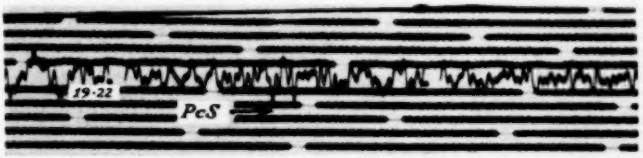


FIG. 5.—1957 April 24 (Aegean Isles). *PcS*—waves recorded at Tolmezzo (Δ = 1733 km).

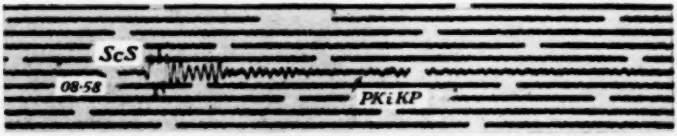


FIG. 9.—1958 June 30 (Southern Sporades). *ScS*—and *PKiKP* recorded at Tolmezzo (Δ = 1630 km).

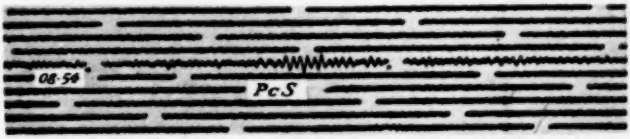


FIG. 10.—1958 June 30 (Southern Sporades). *PcS*—waves recorded at Tolmezzo (Δ = 1630 km).

[*facing page* 144

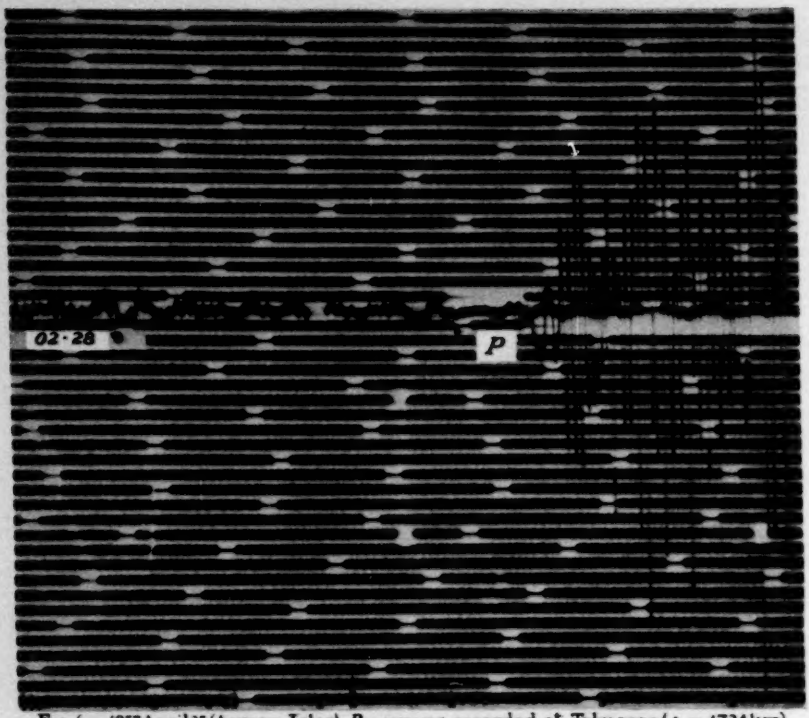


FIG. 6.—1957 April 25 (Aegean Isles). *P*-waves recorded at Tolmezzo (Δ = 1724 km).

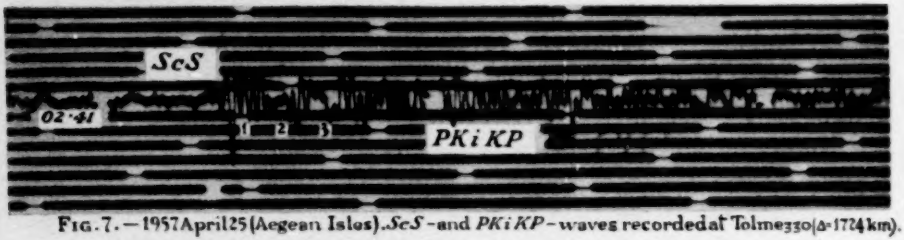


FIG. 7.—1957 April 25 (Aegean Isles). *ScS*-and *PKiKP*-waves recorded at Tolmezzo (Δ=1724 km).

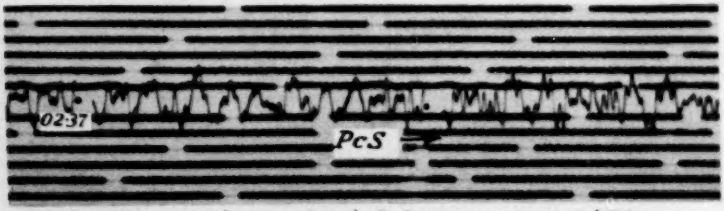


FIG. 8.—1957 April 25 (Aegean Isles). *PcS*-waves recorded at Tolmezzo (Δ = 1724 km)

The second shock shows a focus in the Mohorovičić discontinuity, which is perhaps a little low.

The first shock shows a difference of 4 s plus for the observed value; such difference can be due to a small focal depth (18 km instead of 30 km). The third shock shows a focus in the interior of the crust at the probable depth of 23 km.

In the case of the second earthquake, the depth of 30–33 km is confirmed by a precise registration of *pP* waves 7 s after the beginning (Figure 6).

Remarks

ScS and crustal structure. The Tolmezzo observation of the shock of April 25 (Figure 7) is important not only because it gave an exceptional example of *ScS* waves and of *PKiKP* waves—about which more is said later—but also because it allowed a reconstruction of the Earth's crust in the region of the epicentral zone. It had its focus at the base of the Earth's crust and three arrivals, *1*, *2* and *3*, can be interpreted as travelling in an almost radial direction and reflected at the intermediate layer of the Earth's crust, at the granite layer and at the outer surface of the Earth. Considering the differences $1-ScS$; $2-ScS$; $3-ScS$ and the mean velocity values of 3·8 km/s and 3·7 km/s and 3·6 km/s for the transverse waves in layers concerned, we obtain the following thicknesses for the said layers:

			Thickness
Basalt	4·5 s	*1-ScS*	8·5 km
Basalt + Intermediate	12 s	*2-ScS*	22 km
Basalt + Intermediate + Granite	20 s	*3-ScS*	36 km

Therefore the crust thickness would be of 36 km, made up of 8·5 km of basalt, 13·5 km of intermediate and 14 km of granitic layers. We note that Markus Båth studying the dispersion of the Love waves coming from the same epicentral zone obtained 35 km for the crustal thickness.

In the case of precise *ScS* registrations—and only in this case can the focal depth be determined without laborious calculations—the comparison between the calculated and observed travel-times can give a clear indication of the focal depth.

This is the case of the earthquake of Southern Sporades:

1958 June 30 $H = 08\ 42\ 33$, $\phi = 36°{\cdot}5\,\mathrm{N}$, $\lambda = 27°{\cdot}5\,\mathrm{E}$ U.S.C.G.S.

The Bureau Central of Strasbourg gives the same epicentre, but it shows a depth of about 60 km against the calculation of the U.S.C.G.S. which considers the shock at a normal depth.

The Tolmezzo records give 08 h 58 m 08 s for *ScS* and 1 631 km for Δ.

The corresponding travel-times for *ScS* are 953 s observed and 944·5 s calculated.

The 9 s lag for the calculated value in the case of a focus at the Mohorovičić discontinuity shows a focal depth which is sensibly higher than normal. We may conclude that, on the basis of the *ScS* waves recorded at Tolmezzo, the focal depth

of the Sporades earthquake is about 70 km, confirming the value calculated by the B.C.I.S.

The travel-times calculated with the new formulae for the given examples do practically coincide with those given by Jeffreys & Bullen tables.

The *ScS* waves examples here examined are shown in Figures 4, 7, 9.

5. Records of *PcS* (*ScP*)

The recording of *PcS* (or *ScP*) waves is even more rare than that of *ScS* waves in the case of shocks of normal depth. The observation of such waves is very difficult also in the case of earthquakes at a higher depth (Gutenberg & Richter 1939). Therefore it is evident that there is particular importance in the precise example of *PcS* waves recorded at Tolmezzo in conjunction with the four above reported earthquakes.

The times obtained at Tolmezzo for the three first earthquakes are:

		h	m	s	
1957 April 24	*PcS*	19	22	21·5	$\Delta = 15° 35'$
1957 April 25	*PcS*	02	37	52	$\Delta = 15° 31'$
1957 April 26	*PcS*	06	45	47	$\Delta = 15° 31'$

The observed travel-times are in good agreement with the times calculated by Jeffreys & Bullen for an origin on the Mohorovičić discontinuity (33 km), though the first one agrees rather with a focus around 20 km.

In the case of the fourth shock, that of the Southern Sporades, the record is even clearer (Figure 10) because the *PcS* arrival is more isolated. The time

1958 June 30 *PcS* 08h 54m 36s $\Delta = 14° 40'$,

when referred to the calculated times, gives an earthquake depth about 70 km which agrees with the value obtained by the *ScS*.

6. Recording of *PKiKP*

The real surprise given by the four reported shocks consists in the precise record of longitudinal waves which I have no hesitation in considering as *reflected* waves from the inner core.

In 1938 Gutenberg & Richter confirmed the hypothesis of Miss Lehmann (1936) on the existence of "a discontinuity separating an outer part of the core with a lower wave velocity from an inner core with a greater velocity". They found that *P''* waves can be accounted for as waves refracted through a relatively thin layer (approximately 300 km thick) at a depth of about 5 000 km, in which there is a rapid but probably continuous increase of about 1·2 km/s of velocity with depth (Gutenberg 1951).

Jeffreys (1939a) assumed that true reflection occurs at the inner core, while Gutenberg and Richter had fitted a reversed segment between the two parts of the travel-time curve without supposing reflection. Jeffreys found that on his assumption no agreement is possible, unless the velocity actually decreases with depth in the range of r between about 1 400 and 1 250 km. He pointed out (Jeffreys 1939b page 600) that he had adopted the hypothesis of reflection for convenience, but that

it may well be true from various considerations, among which the observed amplitudes of P' played an important role. In any case, the assumption of such an hypothesis allowed Jeffreys to obtain travel-time values for *PKIKP* waves in agreement with those given by observation.

Gutenberg in the following years made many studies on this subject. He observed (Gutenberg 1957a) that *PKIKP* longitudinal waves with periods of about one second which have passed through the inner core frequently arrive earlier than the corresponding waves with periods of several seconds ($\geqslant 2$). At epicentral distances between about 120° and 135° the difference in arrival between the first short-period waves and the long-period (about 3 s) main impulse may reach nearly 20 s. If this difference is a consequence of dispersion in a transition zone between the outer and inner core the calculated radius of the inner core is greater for short-period than for long-period waves. Therefore, in the core between the outer and the inner core there is no sudden change in material, but a transition from the "liquid" state to one with much higher viscosity. In 1958 he arrived at the same conclusion: "Between about 1 500 and 1 200 km from the Earth's centre in the transition zone from the liquid outer to the probably solid inner core, waves having lengths of the order of 10 km travel faster than longer waves. This is probably caused by a rather rapid increase in viscosity toward the Earth's centre in this transition zone" (Gutenberg 1958a). And further: "Waves with periods of less than 2 s, observed at distances of between 19° and 4° from the caustic of *PKP* are probably *PKIKP* waves with deepest point in or near the transition zone between the liquid outer and the solid inner core; in this zone, short waves travel faster than longer waves" (Gutenberg 1958b). To conclude, in 1959: "Jeffreys assumes that the wave velocity decreases with depth just above the inner core, and increases suddenly at the boundary of the inner core, while Gutenberg and others believe that there is a gradual increase in velocity in a transition zone without a preceding decrease" (Gutenberg 1959).

Gutenberg denied the existence of a reflecting surface as an external boundary of the inner core. Jeffreys admits the existence of this surface only as a working convenience, though he says that it may well be true.

Nevertheless, in 1939 Jeffreys calculated the propagation times of *KiK* rays ("The times of the core waves", second paper), which were not reported in the Jeffreys–Bullen tables. I will here refer only to the results corresponding to the lowest values (in degrees) of:

Δ deg	T s	Δ deg	T s
0	475·9	10	477·4
5	476·2	15	479·3

Such values were obtained considering the hypothesis of a reflecting surface with a radius $r = 0{\cdot}36\,R_1$ (R_1 = core radius = $0{\cdot}548\,R$ (Earth's radius)).

The values calculated for the *PKiKP* waves can be deduced from those for the *KiK* waves, by suitable addition to the travel-times of the *PcP* waves (pages 37–38 of *Seismological Tables* by Jeffreys–Bullen). This has been done for the observations at Tolmezzo, taking into due account the fact that the travel-times for the *PcP* waves are not those referred to the real epicentral distances from Tolmezzo, but those referred to sensibly lower distances; in fact the two prongs of the *PcP* waves are merely the paths of the *PKiKP* waves in the mantle.

The values observed and calculated for the *PKiKP* waves at Tolmezzo are:

	Δ		*PKiKP times* Observed h	m	s	Calculated h	m	s
1957 April 24	15°·6	*ei*	19	26	32	19	26	32
		i			34			
1957 April 25	15°·5		02	42	03	02	42	03
1958 June 30	14°·7	*e*	08	58	50	08	59	00
		ei			54			

We have seen that the reported shocks had different depths, 20, 30 and 75 km respectively, on the basis of the *ScS* waves, and these values are confirmed by the observed times of the *PKiKP* waves. For the Southern Sporades shock (1958 June 30) a depth of at least 85 km is calculated.

In these four shocks the difference *PKiKP*−*ScS* of the observed values resulted as follows:

	PKiKP−*ScS*						
	h	m	s	h	m	s	
1957 April 24	19	26	33	19	25	57	= 36 s
1957 April 25	02	42	03	02	41	25	= 38 s
1957 April 26	06	50	(00)	06	49	23	= (37)
1958 June 30	08	58	50	08	58	08	= 42 s

These figures not only confirm the different focal depths of the examined earthquakes, but also prove that the waves considered are really *PKiKP* waves; in fact when the depth increases, the difference between arrival times of *PKiKP* and *ScS* waves *has also to increase*.

The existence of a reflecting surface, as a boundary of the inner core, is a proof of the solid state of the inner core itself. Besides, it gives a proof that the propagation speed from the outer core to the inner core undergoes a sudden start, as is foreseen in the Jeffreys theory; therefore the transition does not occur gradually as affirmed by Gutenberg.

Following these new results, there is a different explanation of the short-period waves, which Gutenberg indicated by (*a*). These waves precede by a few seconds the (*b*) waves, considered to be *PKIKP* waves. Jeffreys's theory considers the inner core surrounded by a zone of about 150 km thick, which is characterized by a low velocity. Gutenberg on the contrary considers this zone to be one in which the velocity gradually increases above that of the external core. The truth is possibly somewhere between the two hypotheses. The registered *PKiKP* waves (see Figures 4, 7) do not present themselves as a single train of waves; on the contrary, they give place to a long succession of oscillations, with sudden increases. It is possible that the boundary of the inner core shows a solidification into superimposed layers, characterized by increasing speeds, such as those of the Earth's crust. The waves indicated by Gutenberg by (*b*)—the slowest ones—could belong to the very outer layer, and the (*a*) ones could be produced by a refraction phenomenon in the surface limiting the following layer, where the velocity is high enough to cause the (*a*) waves to arrive before the (*b*) waves, in those cases where the distance allows a suitable path for waves refracted in the inner core. Gutenberg refers the two phases (*a*) and (*b*) to an anomalous dispersion, connected with an

increasing velocity in the transition zone. However, the difference in period between the two types of waves is not sufficiently great to admit a definite effect of viscosity (see examples quoted by Gutenberg, page 110 of his last book); furthermore, the distinction between (*a*) and (*b*) waves with reference to their amplitude finds a more plausible explanation as a refraction effect than as a consequence of a gradual increase of the viscosity in the supposed transition zone. The differences between the travel-times of (*a*) and (*b*) waves on the refraction hypothesis, could find an easy explanation in the behaviour of the velocity allowed by Jeffreys in the *F* zone and at the beginning of the *E* zone, where the velocities calculated by Jeffreys are higher than those obtained by Gutenberg.

7. Difficulties in recording *PKiKP* waves

The short period (less than 2 s) and the small amplitude of the *PKiKP* waves allows them to be recorded only by seismographs of short period, with a very large magnification (the "Girlanda" seismograph at Tolmezzo has a magnification of about 9000 for 1 s period waves). Moreover, the most convenient distance for the observation of these waves are those under 20°. In fact, in the case of greater distances, the *PKiKP* waves ultimately fall little by little from the phase of surface waves to the phase of reflected longitudinal body waves. It is also to be pointed out that—with regard to the quoted examples—the underground at Tolmezzo station is particularly suited to the recording of short-period waves. The thick alluvial layer where the instrument is placed amplifies the short periods: that is to say that the place has evidently that "Untergrundsfaktor" of which Gutenberg spoke in 1929 (Gutenberg 1929) and about which he has recently written (Gutenberg 1957).

Istituto Nazionale di Geofisica,
Roma:
1960 *August.*

References

Båth, M. & Vogel, A., 1957. Continental dispersion of seismic surface waves. *Geof. pura e applicata*, **38**, 10–18.

Caloi, P. & Peronaci, F., 1949. Il batisismo del 28 agosto 1946 e la profodità del nucleo terrestre. *Annali di Geofisica*, II.

Gutenberg, B., 1929. Theorie der Erdbebenwellen; Beobachtungen; Bodenunruhe. *Handbuch der Geophysik*, Band IV, Lief. 1, 193 (Berlin).

Gutenberg, B., 1951. *PKKP*, *P'P'*, and the Earth's core. *Trans. Amer. Geophys. Un.*, **32**, n. 3.

Gutenberg, B., 1957a. The boundary of the Earth's inner core. *Trans. Amer. Geophys. Un.*, **38**, 753.

Gutenberg, B., 1957b. Effects of ground on earthquake motion. *Bull. Seismol. Soc. Amer.*, **47**, 221–250.

Gutenberg, B., 1958a. Waves velocity in the Earth's core. *Bull. Seismol. Soc. Amer.*, **48**, 301.

Gutenberg, B., 1958b. Caustics produced by waves through the Earth's core. *Geophys. J.*, **1**, 238.

Gutenberg, B., 1959. *Physics of the Earth's Interior* (Academic Press), 111.

Gutenberg, B. & Richter, C. F., 1938. *P'* and the Earth's core. *Mon. Not. R. Astr. Soc. Geophys. Suppl.*, **4**, 363–372.

Gutenberg, B. & Richter, C. F., 1939. On seismic waves. Fourth paper, *Beitr. Geophys.* (*Gerlands*), **54**, 105.

Jeffreys, H., 1939a. The times of the core waves. *Mon. Not. R. Astr. Soc. Geophys. Suppl.*, **4**, 548–561, 594–615.

Jeffreys, H., 1939b. *The Earth* (Cambridge: University Press).

Lehmann, I., 1936. *P'*. *Pub. Bureau Central Séismol. Internat., Trav. Sci. Sér. A.* **14**, 87–115.

Solution of an n-layer problem by a seismic reflection method

Fritz Gassmann

"A brief reflection solves the mystery."—*William Stubbs.*

Summary

The treatment of the n-layer problem presented in this paper is based on the assumption that the subsoil is bounded by a plane surface and that plane interfaces of arbitrary position subdivide the subsoil into homogeneous isotropic solid layers.

In Section 1 recursion formulae are derived for wave-rays which are reflected by one interface and refracted by the others.

It is shown in Section 2 that a parametric representation of the travel-time function can be found for an arbitrary position of the point source.

Section 3 deals with the computation of the travel-time for a given position of a seismometer. The method is based on successive approximations using a Taylor expansion of the travel-time.

In Sections 4 and 5 a method is presented to compute the front velocities and the positions of the interfaces in the subsoil on the basis of suitable seismic reflection measurements.

1. Definitions and basic equations

It is assumed that the subsoil is bounded by a plane surface (S_1) and subdivided by plane interfaces (S_2), (S_3), ... into homogeneous isotropic solid layers. The positions of the interfaces are defined by a chain of vectors $\mathbf{h}_1, \mathbf{h}_2, \ldots$ (Figure 1), $\mathbf{h}_i = O_iO_{i+1}$ which are perpendicular to (S_i), $i = 1, 2, \ldots$. Weber (1959) has treated the case of a plane polygon $\mathbf{h}_1, \mathbf{h}_2, \ldots$. In the present paper the methods of M. Weber are generalized and applied to an arbitrary three-dimensional polygon. Hence Figure 1 represents only a schematic representation of the actual three-dimensional reality. Let Q be a point source situated either in the plane (S_k) or in the kth layer, i.e. between (S_k) and (S_{k+1}), $O_kQ = \mathbf{a}$. Consider a P-wave front with origin in Q which is reflected at (S_n), $n > k$.

$$QQ_{k+1} \ldots Q_{n-1} P_nP_{n-1} \ldots P_2P_1 \; (Q_n = P_n)$$

is a corresponding ray (see Figure 1, where $n = 4$, $Q = Q_1 = O_1$, i.e. $k = 1$, $\mathbf{a} = 0$). C_i is the front velocity in the ith layer, $\lambda_i = c_i/c_{i-1}$, $\mathbf{b} = \mathbf{QQ}_{k+1}$, $\mathbf{s}_i = \mathbf{Q}_i\mathbf{Q}_{i+1}$, $\mathbf{r}_i = \mathbf{P}_{i+1}\mathbf{P}_i$. The absolute value of a vector is denoted by the same letter as the vector, but not in bold type, e.g. $s_i = |\mathbf{s}_i|$. $\boldsymbol{\epsilon}_i = \mathbf{h}_i/h_i$,

$\mathbf{e}_k = \mathbf{b}/b$, $\mathbf{e}_i = \mathbf{s}_i/s_i$, $\mathbf{d}_i = \mathbf{r}_i/r_i$ are unit vectors. At the points P_i and Q_i, the ray is refracted and the law of refraction can be expressed as follows:

$$\mathbf{d}_i = \lambda_i \mathbf{d}_{i-1} - [\lambda_i \mathbf{e}_i \mathbf{d}_{i-1} + \{1 - \lambda_i^2 (\mathbf{e}_i \wedge \mathbf{d}_{i-1})^2\}^{\frac{1}{2}}]\mathbf{e}_i, \tag{1}$$

$$\mathbf{e}_i = \lambda_i \mathbf{e}_{i-1} - [\lambda_i \mathbf{e}_i \mathbf{e}_{i-1} - \{1 - \lambda_i^2 (\mathbf{e}_i \wedge \mathbf{e}_{i-1})^2\}^{\frac{1}{2}}]\mathbf{e}_i, \tag{2}$$

or

$$\mathbf{e}_{i-1} = \frac{1}{\lambda_i}\mathbf{e}_i - \left[\frac{1}{\lambda_i}\mathbf{e}_i \mathbf{e}_i - \left\{1 - \frac{1}{\lambda_i^2}(\mathbf{e}_i \wedge \mathbf{e}_i)^2\right\}^{\frac{1}{2}}\right]\mathbf{e}_i. \tag{3}$$

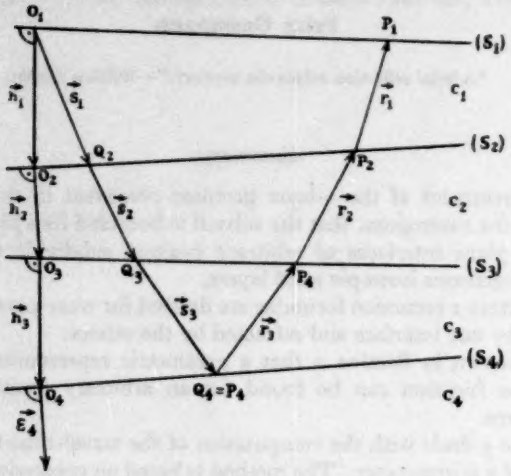


FIG. 1.

At P_n, the law of reflection is

$$\mathbf{e}_{n-1} = \mathbf{d}_{n-1} - 2(\mathbf{d}_{n-1}, \boldsymbol{\epsilon}_n) \cdot \boldsymbol{\epsilon}_n. \tag{4}$$

If the position of Q and the vectors $\mathbf{h}_i$, $\mathbf{e}_i$ and $\mathbf{d}_i$ are given, the distances b, s_i, r_i and the vectors $\mathbf{y}_i = O_iQ_i$, $\mathbf{x}_i = O_iP_i$ can be computed. For this purpose the conditions $\mathbf{e}_i\mathbf{e}_{i+1} > 0$ and $\mathbf{d}_i\mathbf{e}_i < 0$ are assumed to be valid and use is made of the following equations:

$$\boldsymbol{\epsilon}_i \mathbf{y}_i = 0, \qquad \boldsymbol{\epsilon}_i \mathbf{x}_i = 0.$$

$$\mathbf{y}_{k+1} = \mathbf{a} + b\mathbf{e}_k - \mathbf{h}_k. \tag{5}$$

$$b = \frac{(\mathbf{h}_k - \mathbf{a})\boldsymbol{\epsilon}_{k+1}}{\mathbf{e}_k \boldsymbol{\epsilon}_{k+1}}. \tag{6}$$

$$\mathbf{y}_{i+1} = \mathbf{y}_i + s_i \mathbf{e}_i - \mathbf{h}_i \quad \text{for} \quad k < i < n. \tag{7}$$

$$s_i = \frac{(\mathbf{h}_i - \mathbf{y}_i)\boldsymbol{\epsilon}_{i+1}}{\mathbf{e}_i \boldsymbol{\epsilon}_{i+1}} \quad \text{for} \quad k < i < n. \tag{8}$$

$$\mathbf{x}_n = \mathbf{y}_n. \tag{9}$$

$$\mathbf{x}_i = \mathbf{h}_i + r_i \mathbf{d}_i + \mathbf{x}_{i+1}. \tag{10}$$

$$r_i = -\frac{(\mathbf{h}_i + \mathbf{x}_{i+1})\boldsymbol{\epsilon}_i}{\mathbf{d}_i \boldsymbol{\epsilon}_i}. \tag{11}$$

With the aid of these equations, the quantities b, $\mathbf{y}_{k+1}$, s_{k+1}, $\mathbf{y}_{k+2}$, ... s_{n-1}, $\mathbf{y}_n$, $\mathbf{x}_n$ r_{n-1}, $\mathbf{x}_{n-1}$, r_{n-2}, ... r_1, $\mathbf{x}_1$ are obtained step by step.

2. The travel-time function

Let O_1 be the origin of the Cartesian coordinate system, ξ_1, ξ_2, ξ_3. The unit vectors η_1, η_2, η_3 point in the positive directions of the coordinates. (S_1) is the plane $\xi_3 = 0$ and the $+\xi_3$-axis points downwards. Therefore $\epsilon_1 = \eta_3$. Consider a wave front with origin in O_1 which is reflected at the plane (S_n). The travel-time of this wave front from O_1 to a point P in the first layer is a function $\tau_n(\xi_1, \xi_2, \xi_3)$ of the coordinates of P.

$$\frac{\partial \tau_n}{\partial \xi_1} = p_{n1}, \quad \frac{\partial \tau_n}{\partial \xi_2} = p_{n2} \quad \text{and} \quad \frac{\partial \tau_n}{\partial \xi_3} = p_{n3}$$

are the components of grad τ_n and

$$c_1 \mathbf{d}_1 = \frac{\operatorname{grad} \tau_n}{(\operatorname{grad} \tau_n)^2}$$

is the front velocity at P. Therefore

$$p_{n1}^2 + p_{n2}^2 + p_{n3}^2 = \frac{1}{c_1^2}$$

and

$$\mathbf{d}_1 = c_1 p_{n1} \eta_1 + c_1 p_{n2} \eta_2 - [1 - c_1^2 (p_{n1}^2 + p_{n2}^2)]^{\frac{1}{2}} \eta_3. \tag{12}$$

If P is a point P_1 of the plane (S_1), the travel-time

$$\tau_n = \frac{b}{c_k} + \sum_{i=k+1}^{n-1} \frac{s_i}{c_i} + \sum_{i=1}^{n-1} \frac{r_i}{c_i} \tag{13}$$

is a function of the position of P_1, i.e. of

$$\mathbf{x}_1 = \mathbf{O}_1 \mathbf{P}_1 = \xi_1 \eta_1 + \xi_2 \eta_2. \tag{14}$$

Let us assume that $\mathbf{h}_1$, $\mathbf{h}_2$, ... $\mathbf{h}_{n-1}$, ϵ_n, c_1, ... c_{n-1} and the position of Q are known and that the values of p_{n1} and p_{n2} are given. By virtue of Equations (12), (1), (4) and (3), the unit vectors $\mathbf{d}_1, \mathbf{d}_2, \ldots \mathbf{d}_{n-1}, \mathbf{e}_{n-1}, \mathbf{e}_{n-2}, \ldots \mathbf{e}_k$ can be computed step by step. Then b, s_i, r_i and $\mathbf{x}_1$ are obtained using the methods of Section 1 and τ_n is determined by Equation (13). In this manner a parametric representation

$$\mathbf{x}_1 = \mathbf{x}_1(p_{n1}, p_{n2}), \qquad \tau_n = \tau_n(p_{n1}, p_{n2}) \tag{15}$$

of the travel-time function $\tau_n(\mathbf{x}_1)$ is obtained.

3. A Taylor expansion for the travel-time

First some derivatives with respect to p_{n1} and p_{n2} are needed. From (12), one gets

$$\frac{\partial \mathbf{d}_1}{\partial p_{nj}} = c_1 \eta_j + \frac{c_1^2 p_{nj}}{[1 - c_1^2 (p_{n1}^2 + p_{n2}^2)]^{\frac{1}{2}}} \cdot \eta_3. \tag{16}$$

The index j is either 1 or 2.

Differentiation of (1) gives

$$\frac{\partial \mathbf{d}_i}{\partial p_{nj}} = \lambda_i \frac{\partial \mathbf{d}_{i-1}}{\partial p_{nj}} - \left\{ \lambda_i \mathbf{e}_i \frac{\partial \mathbf{d}_{i-1}}{\partial p_{nj}} - \frac{\lambda_i^2 (\mathbf{e}_i \wedge \mathbf{d}_{i-1})[\mathbf{e}_i \wedge (\partial \mathbf{d}_{i-1}/\partial p_{nj})]}{\sqrt{[1-\lambda_i^2(\mathbf{e}_i \wedge \mathbf{d}_{i-1})^2]}} \right\} \cdot \mathbf{e}_i. \tag{17}$$

From equations (4) and (3) the expressions for $\partial \mathbf{e}_{n-1}/\partial p_{nj}$ and $\partial \mathbf{e}_{i-1}/\partial p_{nj}$ are obtained. Thus all these derivatives are considered to be known functions of p_{n1} and p_{n2}. From the corresponding equations of Section 2, the derivatives $\partial \mathbf{x}_1/\partial p_{nj}$ can be computed in the same way as $\mathbf{x}_1(p_{n1}, p_{n2})$ was obtained in Section 2; i.e. $\partial \xi_1/\partial p_{nj}$ and $\partial \xi_2/\partial p_{nj}$,

$$D = \frac{\partial(\xi_1, \xi_2)}{\partial(p_{n1}, p_{n2})} = \frac{\partial \xi_1}{\partial p_{n1}} \cdot \frac{\partial \xi_2}{\partial p_{n2}} - \frac{\partial \xi_1}{\partial p_{n2}} \cdot \frac{\partial \xi_2}{\partial p_{n1}}, \tag{18}$$

as well as

$$\begin{aligned} \frac{\partial^2 \tau_n}{\partial \xi_1^2} &= \frac{\partial p_{n1}}{\partial \xi_1} = \frac{1}{D} \frac{\partial \xi_2}{\partial p_{n2}} \\ \frac{\partial^2 \tau_n}{\partial \xi_1 \partial \xi_2} &= \frac{\partial p_{n1}}{\partial \xi_2} = \frac{\partial p_{n2}}{\partial \xi_1} = -\frac{1}{D} \frac{\partial \xi_1}{\partial p_{n2}} \\ \frac{\partial^2 \tau_n}{\partial \xi_2^2} &= \frac{\partial p_{n2}}{\partial \xi_2} = \frac{1}{D} \frac{\partial \xi_1}{\partial p_{n1}} \end{aligned} \tag{19}$$

are known functions of p_{n1} and p_{n2}.

Now let us indicate by p_{n1}^0 and p_{n2}^0 certain fixed values of p_{n1} and p_{n2}. (The corresponding fixed values of other quantities shall be designated by the upper index 0.) According to (14) and (15),

$$\mathbf{x}_1^0 = \mathbf{x}_1(p_{n1}^0, p_{n2}^0) = \xi_1^0 \eta_1 + \xi_2^0 \eta_2 \tag{20}$$

and the Taylor expansion of τ_n in the neighbourhood of $\mathbf{x}_1^0$ gives the approximation

$$\begin{aligned} \tau_n \cong \tau_n^0 + p_{n1}^0(\xi_1 - \xi_1^0) + p_{n2}^0(\xi_2 - \xi_2^0) + \frac{1}{2}\left(\frac{\partial p_{n1}}{\partial \xi_1}\right)^0 (\xi_1 - \xi_1^0) + \\ + \left(\frac{\partial p_{n1}}{\partial \xi_2}\right)^0 (\xi_1 - \xi_1^0)(\xi_2 - \xi_2^0) + \frac{1}{2}\left(\frac{\partial p_{n2}}{\partial \xi_2}\right)^0 (\xi_2 - \xi_2^0)^2 \end{aligned} \tag{21}$$

and by differentiation

$$p_{nj} \cong p_{nj}^0 + \left(\frac{\partial p_{nj}}{\partial \xi_1}\right)^0 (\xi_1 - \xi_1^0) + \left(\frac{\partial p_{nj}}{\partial \xi_2}\right)^0 (\xi_2 - \xi_2^0). \tag{22}$$

As has been shown by Weber the formulae (21) and (22) can be used to compute $\tau_n(\mathbf{x}_1)$ by successive approximations for a given value of $\mathbf{x}_1$. For the corresponding (unknown) quantities p_{nj}, approximate values p_{nj}^0 can be estimated. By virtue of equations (20) and (15), the values of ξ_1^0, ξ_2^0 and τ_n^0 are obtained and, as has been shown above, the expression (21) can be computed. This procedure gives a first approximation for τ_n. If higher approximations are desired, the procedure is repeated by substituting for p_{n1}^0 and p_{n2}^0 values obtained from (22).

4. Solution of the 2-layer problem

The velocities c_i and the positions of the interfaces (S_i), i.e. the vectors $\mathbf{h}_i$ can be determined by seismic reflexion measurements at the surface (S_1), if they are not known. First, the point O_1 is chosen as shotpoint, i.e. $k = 1$ and $\mathbf{a} = 0$. Using a suitable layout of seismometers at P_1 and in its neighbourhood, the travel-times τ_n are measured and the quantities $p_{n1} = \partial\tau_n/\partial\xi_1$ and $p_{n2} = \partial\tau_n/\partial\xi_2$ $(n = 2, 3, \ldots)$ are determined. Second, P_1 is chosen as shotpoint and the seismometers are located at O_1 and in its neighbourhood. With this arrangement, the travel-times τ_n are measured and the quantities q_{n1} and q_{n2}, which correspond to p_{n1} and p_{n2}, are determined. q_{n1} and q_{n2} are defined by an equation, corresponding to equation (12), i.e.

$$\mathbf{e}_1 = c_1 q_{n1}\eta_1 + c_1 q_{n2}\eta_2 + [1 - c_1^2(q_{n1}^2 + q_{n2}^2)]^{\frac{1}{2}}\eta_3. \tag{23}$$

In order to determine the unknown quantities c_i and $\mathbf{h}_i$, one has to proceed step by step. First the 2-layer problem must be solved. Its solution requires (a) the determination of the quantities $\tau_2, p_{21}, p_{22}, q_{21}, q_{22}$ for a given $\mathbf{x}_1$, and (b) the computation of $c_1, \mathbf{e}_2$ and h_1. The following equations hold for the ray $O_1P_2P_1$ which is reflected at (S_2):

$$\mathbf{x}_1 = s_1\mathbf{e}_1 + r_1\mathbf{d}_1, \qquad \mathbf{x}_1(\mathbf{e}_1 + \mathbf{d}_1) = (s_1 + r_1)(1 + \mathbf{e}_1\mathbf{d}_1),$$

$$s_1 + r_1 = c_1\tau_2, \qquad \theta_1 = \frac{\mathbf{x}_1(\mathbf{e}_1 + \mathbf{d}_1)}{c_1\tau_2}.$$

By virtue of equations (14), (23) and (12), (for $n = 2$) the number

$$\theta_1 = \frac{\xi_1(q_{21} + p_{21}) + \xi_2(q_{22} + p_{22})}{\tau_2} \tag{24}$$

is known. It follows

$$\mathbf{e}_1\mathbf{d}_1 + 1 - \theta_1 = 0. \tag{25}$$

After substitution of the expressions (23) and (12) into equation (25) and after elimination of the square roots, an equation of second degree for c_1^2 results. From this equation the desired value of c_1 is obtained. By virtue of (12) and (23), $\mathbf{d}_1$ and $\mathbf{e}_1$ are computed. For $n = 2$, equation (4) becomes

$$\mathbf{e}_1 = \mathbf{d}_1 - 2(\mathbf{d}_1\boldsymbol{\epsilon}_2)\,.\,\boldsymbol{\epsilon}_2.$$

Therefore

$$\mathbf{d}_1\mathbf{e}_1 = 1 - 2(\mathbf{d}_1\boldsymbol{\epsilon}_2)^2, \qquad \mathbf{d}_1\boldsymbol{\epsilon}_2 = -\sqrt{\frac{1 - \mathbf{d}_1\mathbf{e}_1}{2}},$$

$$\boldsymbol{\epsilon}_2 = \frac{\mathbf{e}_1 - \mathbf{d}_1}{[2(1 - \mathbf{d}_1\mathbf{e}_1)]^{\frac{1}{2}}}. \tag{26}$$

For the ray $O_1P_2P_1$, reflected at (S_2),

$$s_1\mathbf{e}_1 = \mathbf{h}_1 + \mathbf{y}_2, \qquad r_1\mathbf{d}_1 = \mathbf{x}_1 - \mathbf{h}_1 - \mathbf{y}_2.$$

Multiplication with $\mathbf{e}_1$ and $\mathbf{d}_1$ respectively and addition gives

$$s_1 + r_1 = c_1\tau_2 = \mathbf{h}_1(\mathbf{e}_1 - \mathbf{d}_1) + \mathbf{x}_1\mathbf{d}_1$$
$$= \boldsymbol{\epsilon}_1 h_1(\mathbf{e}_1 - \mathbf{d}_1) + \mathbf{x}_1\mathbf{d}_1, \qquad \text{since } \mathbf{y}_2(\mathbf{e}_1 - \mathbf{d}_1) = 0 \text{ (see (26))}.$$

Thus

$$h_1 = \frac{c_1\tau_2 - \mathbf{x}_1\mathbf{d}_1}{\epsilon_1(\mathbf{e}_1 - \mathbf{d}_1)}. \tag{27}$$

Equations (25), (26) and (27) represent the solution of the 2-layer problem.

5. Solution of the n-layer problem

In the following a method is presented which gives a solution of the n-layer problem, provided the solution of the $n-1$-layer problem is known, i.e. $c_1, c_2 \ldots c_{n-2}$, $\mathbf{h}_1, \mathbf{h}_2, \ldots \mathbf{h}_{n-2}, \epsilon_{n-1}$ are known and c_{n-1}, ϵ_n and h_{n-1} must be determined. From the kind of measurements described in Section 4, the quantities $\tau_n, p_{n1}, p_{n2}, q_{n1}$ and q_{n2} are determined. By virtue of (23), (2), (12) and (1), the quantities $\mathbf{e}_1, \mathbf{e}_2, \ldots \mathbf{e}_{n-2}$; $\mathbf{d}_1, \mathbf{d}_2, \ldots \mathbf{d}_{n-2}$ are computed. Since $\mathbf{y}_1 = 0$, and $\mathbf{x}_1$ is known, $s_1, \mathbf{y}_2, s_2, \mathbf{y}_3, s_3, \ldots \mathbf{y}_{n-2}, s_{n-2}, \mathbf{y}_{n-1}; r_1, \mathbf{x}_2, r_2, \mathbf{x}_3, r_3, \ldots \mathbf{x}_{n-2}, r_{n-2}, \mathbf{x}_{n-1}$ are obtained on the basis of equations (8), (7), (28) and (10), where

$$r_i = \frac{(\mathbf{x}_i - \mathbf{h}_i)\epsilon_{i+1}}{\mathbf{d}_i\epsilon_{i+1}} \tag{28}$$

similar to equation (8).

Now τ_n^* can be computed:

$$\tau_n^* = \tau_n - \sum_{i=1}^{n-2} \frac{r_i + s_i}{c_i}. \tag{29}$$

From

$$\mathbf{x}_{n-1} - \mathbf{y}_{n-1} = s_{n-1}\mathbf{e}_{n-1} + r_{n-1}\mathbf{d}_{n-1},$$

one finds

$$\begin{aligned}(\mathbf{x}_{n-1} - \mathbf{y}_{n-1})(\mathbf{e}_{n-1} + \mathbf{d}_{n-1}) &= (s_{n-1} + r_{n-1})(1 + \mathbf{e}_{n-1}\mathbf{d}_{n-1}) \\ &= c_{n-1}\tau_n^*(1 + \mathbf{e}_{n-1}\mathbf{d}_{n-1}).\end{aligned}$$

If we set in equations (1) and (2) $i = n-1$, the sum

$$\mathbf{e}_{n-1} + \mathbf{d}_{n-1} = \frac{c_{n-1}}{c_{n-2}}(\mathbf{e}_{n-2} + \mathbf{d}_{n-2}) - S^*\epsilon_{n-1}$$

is obtained. S^* is a scalar. Since

$$(\mathbf{x}_{n-1} - \mathbf{y}_{n-1})\epsilon_{n-1} = 0,$$

one obtains

$$(\mathbf{x}_{n-1} - \mathbf{y}_{n-1})(\mathbf{e}_{n-1} + \mathbf{d}_{n-1}) = \frac{c_{n-1}}{c_{n-2}}(\mathbf{x}_{n-1} - \mathbf{y}_{n-1})(\mathbf{e}_{n-2} + \mathbf{d}_{n-2}) \tag{30}$$

and

$$\mathbf{e}_{n-1}\mathbf{d}_{n-1} + 1 - \theta_{n-1} = 0. \tag{31}$$

In this equation

$$\theta_{n-1} = \frac{1}{c_{n-2}\tau_n^*}(\mathbf{x}_{n-1} - \mathbf{y}_{n-1})(\mathbf{e}_{n-2} + \mathbf{d}_{n-2}) \tag{32}$$

is a known quantity. If the expressions (1) and (2) for $i = n-1$ are substituted into (31), an equation of second degree for λ_{n-1}^2 and hence the value of c_{n-1} is obtained. As in Section 4 one finds

$$\epsilon_n = \frac{e_{n-1} - d_{n-1}}{[2(1 - d_{n-1}e_{n-1})]^{\frac{1}{2}}} \tag{33}$$

and, on the basis of equations (7) and (10) for $i = n-1$ one finds

$$h_{n-1} = \frac{c_{n-1}\tau_n^* - x_{n-1}d_{n-1} + y_{n-1}e_{n-1}}{\epsilon_{n-1}(e_{n-1} - d_{n-1})}. \tag{34}$$

This represents the solution of the n-layer problem.

Acknowledgment

I am greatly indebted to my son-in-law, Professor Rolf M. Steffen, for correcting the English.

Institut für Geophysik,
Eidg. Technische Hochschule,
Zürich 6, Switzerland.

1960 *July.*

Reference

Weber, M., 1959. Das ebene n-Schichtenproblem der Reflexionsseismik mit konstanten Frontgeschwindigkeiten, beliebigen Neigungen und beliebiger Lage der Quelle. *Geofisica pura e applicata*, **43** (II), 12–22; and Mitteilungen aus dem Institut für Geophysik No. 36.

III. Seismology — Surface Waves

PAGE

L. KNOPOFF, Green's function for eigenvalue problems and the inversion of Love wave dispersion data 161

E. R. LAPWOOD, The transmission of a Rayleigh pulse round a corner 174

U. HOCHSTRASSER & R. STONELEY, The transmission of Rayleigh waves across an ocean floor with two surface layers 197

R. L. KOVACH & F. PRESS, Rayleigh dispersion and crustal structure in the Eastern Pacific and Indian Oceans 202

Green's Function for Eigenvalue Problems and the Inversion of Love Wave Dispersion Data*

Leon Knopoff

"Bella figlia dell'amore."—*Verdi's Rigoletto, Act IV Quartet*

Summary

The change in the nature of the dispersion curve for Love waves has been determined for the case of a delta function inhomogeneity introduced at depth. The unperturbed medium is arbitrarily inhomogeneous in the depth dimension. It is shown that the perturbation in the dispersion curves due to a density inhomogeneity are decoupled from those due to a modulus inhomogeneity. Hence it is proved that any dispersion curve can correspond to an infinity of possible distributions of density and modulus; the relation between the possible solutions is given by a linear integral equation.

The inversion of the integral equations is carried out for the case when the starting values are those of the usual homogeneous layer overlying a homogeneous half-space. Hence from a given dispersion curve, the infinity of solutions are obtained as perturbations upon the assumed starting structure. The limitations of the inversion are discussed.

1. Introduction

In the theory of the scattering of wave motions by inhomogeneities two alternative procedures are available for the determination of the scattered wave functions. In one of these, the Born approximation, it is assumed the inhomogeneity is small compared with the wavelength. The inhomogeneity is then assumed to be excited as a virtual source by the incident wave; the strength of the excitation depends upon the strength of the wave function present at the inhomogeneity and the degree of inhomogeneity of the obstacle. The solution is usually expressed by an iteration since the wave function at the inhomogeneity is often difficult to express exactly; this is often a consequence of incompatible geometries of the incident wave and the scattering centre.

In the other scheme, originally rediscovered by Jeffreys (1924) (cf. Jeffreys & Jeffreys 1956) and called the JWKB method by atomic physicists, the wavelengths are assumed to be small compared with the geometrical extent of the disturbance. The solutions are written in integral form as wave functions of slowly varying amplitude but with large variations in phase. The exact solution is an integral

* Publication No. 188, Institute of Geophysics, University of California.

equation; again the exact solution is difficult to obtain but may in turn be approximated by an iteration. A summary of these two methods may be found in many texts (cf. Mott & Massey 1933, Morse & Feshbach 1953, etc.).

Scattering theory, as usually expressed, deals with perturbations of the eigenfunctions in the region outside the scattering centre. Measurements of the scattered waves are observed in the region outside the scattering centre. An inversion procedure is necessary to determine the nature of the scattering potential that gives rise to the observed scattered waves; this is done in many atomic and nuclear scattering problems for both the long and short wavelength approximations.

In the propagation of elastic surface waves, one is often interested in the eigenvalue problem rather than in the eigenfunction problem. The zeros of a homogeneous set of equations determine the eigenvalues for the surface waves corresponding to the geometry. A number of geometries have been investigated in this way (summarized in Ewing, Jardetzky & Press 1957, Chapters 4, 5 and 7).

Of perhaps more direct interest is the inversion problem for surface wave eigenvalues. From observations of the dispersion of surface waves as a function of the frequency it is desired to describe the corresponding Earth structure. It may be anticipated that existing solutions may be treated by perturbation methods to determine the effect of small variations in structure upon the dispersion curves.

The case of wavelengths that are short when compared with the dimensions of the inhomogeneity has already been treated by Takahashi (1955, 1957) by an application of the JWKB method. The procedure is applicable only to short wavelength surface waves and therefore can be applied with assurance to the determination of near surface Earth structures. Another difficulty is encountered in Takahashi's work; no recognition is made of the lack of uniqueness in surface wave inversion. It will be shown later in this paper that the elastic parameters and the density parameter are decoupled and hence that perturbations in surface wave dispersion can be independently associated with either of these two parameters; thus a further assumption regarding the variation of these parameters must be obtained.

Of interest, at the present time, is the variation in velocity and density with depth in the mantle of the Earth. It is of interest to see whether there are significant differences in the velocities determined by inversion and the velocity–depth functions obtained by Jeffreys and Bullen and Gutenberg from travel-time data; a similar comparison should be made between the densities determined by dispersion curve inversion and the density–depth functions obtained by Bullen. This is especially so in view of the consideration given in recent years to the possibility of low velocity regions in the mantle.

The numerous measurements of surface wave dispersion data obtained from long-period seismograph records suggest that an approach, similar to that of the Born approximation in scattering theory, be used to develop an inversion procedure for the surface wave eigenvalue problem. Such a procedure is given in this paper for Love waves; the corresponding result for Rayleigh waves will appear in a forthcoming paper.

2. Perturbation theory: upper layer

Consider an inhomogeneous Earth with elastic modulus μ and density ρ varying only with the Cartesian coordinate z. The differential equation satisfied by the

Fourier transform of the particle displacement u in the SH polarization is

$$\mu\frac{\partial^2 u}{\partial x^2}+\frac{\partial}{\partial z}\mu\frac{\partial u}{\partial z}+\rho\omega^2 u = 0. \tag{1}$$

We assume that surface waves travelling in the x-direction exist. The wave number is taken to be $k(\omega)$. The solution is thus

$$u = F(z)e^{ikx}.$$

Substituting this function in (1) we obtain

$$F''+(\mu'/\mu)F'+F(\rho\omega^2/\mu-k^2) = 0 \tag{2}$$

where primes denote differentiation with respect to z. Let the surface of the Earth be $z = 0$. At the frequency ω let there be a certain depth $d(\omega)$ below which the motion decays exponentially with depth. Above $d(\omega)$, $F(z)$ has two parts corresponding to reflection from the surface d. We write, in this region, $F(z) = Af_+(z)+Bf_-(z)$. Below the depth d, $F(z) = Cg_+(z)$. The $\pm$ subscripts refer to the signs of the exponents in the case of a homogeneous layer overlying a homogeneous half-space; in this special case

$$f_\pm = e^{\mp\eta z}$$

$$g_+ = e^{-\eta z}$$

where $\eta^2 = k^2-\rho\omega^2/\mu$. The values of η may be different in the two expressions corresponding to the different properties of the two regions; η^2 in the lower region is positive. In the general case the $\pm$ signs refer to two solutions to the differential equation chosen so that one vanishes at $+\infty$, the other at $-\infty$ if f is real, and correspond to waves travelling in the positive or negative z directions if it is complex or imaginary. In the more general case $f_\pm$ are not necessarily sinusoids or hyperbolic functions but may be more complicated functions. The geometry is shown in Figure 1.

Between the depths $z = b$ and $z = c$ a homogeneous layer is introduced, such that the solutions in the region are $h_\pm(z)e^{ikx}$. The eigenvalue problem for Love waves is resolved to the determination of the zeros of the equation

$$\begin{vmatrix} f'_{+a} & f'_{-a} & 0 & 0 & 0 & 0 & 0 \\ f_{+b} & f_{-b} & h_{+b} & h_{-b} & 0 & 0 & 0 \\ f'_{+b} & f'_{-b} & \psi_b h'_{+b} & \psi_b h'_{-b} & 0 & 0 & 0 \\ 0 & 0 & h_{+c} & h_{-c} & f_{+c} & f_{-c} & 0 \\ 0 & 0 & \psi_c h'_{+c} & \psi_c h'_{-c} & f'_{+c} & f'_{-c} & 0 \\ 0 & 0 & 0 & 0 & f_{+d} & f_{-d} & g_{+d} \\ 0 & 0 & 0 & 0 & f'_{+d} & f'_{-d} & \Omega g'_{+d} \end{vmatrix} = 0 \tag{3}$$

where f'_{+a} implies $(df_+/dz)_{z=a}$, etc. The symbol ψ_b is the ratio of the shear modulus of the region below $z = b$ to that just above $z = b$. ψ_c is the ratio of the shear modulus of the region above $z = c$ to that just below $z = c$. Ω is the ratio of the shear modulus below and above $z = d$ in the immediate vicinity of this boundary.

The upper surface is assumed to be free of shear stress. We further note that the equation

$$L(k) = \begin{vmatrix} f'_{+a} & f'_{-a} & 0 \\ f_{+d} & f_{-d} & g_{+d} \\ f'_{+d} & f'_{-d} & \Omega g'_{+d} \end{vmatrix} = 0 \tag{4}$$

is the ordinary equation for Love waves in the absence of the perturbing layer $b \leqslant z \leqslant c$; $L(k)$ is observed to be a constituent part of equation (3).

We define

$$H_1 = \begin{vmatrix} h'_{+b} & h'_{-b} \\ h_{+c} & h_{-c} \end{vmatrix} \qquad H_2 = \begin{vmatrix} h'_{+b} & h'_{-b} \\ h'_{+c} & h'_{-c} \end{vmatrix}$$

$$H_3 = \begin{vmatrix} h_{+b} & h_{-b} \\ h_{+c} & h_{-c} \end{vmatrix} \qquad H_4 = \begin{vmatrix} h_{+b} & h_{-b} \\ h'_{+c} & h'_{-c} \end{vmatrix}.$$

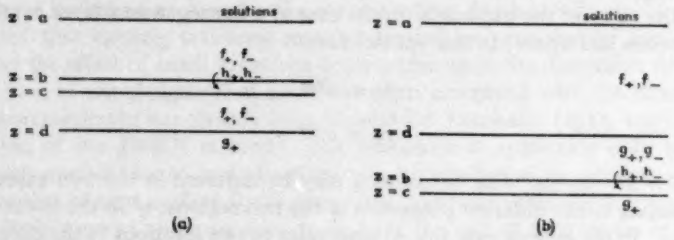


FIG. 1.—(a) Geometry for case of inhomogeneous layer $a < b < z < c < d$.
(b) Geometry for case of inhomogeneous layer $a < d < b < z < c$.

Taking care to expand (3) in such a way as to preserve the properties of (4), we obtain

$$f'_{+a}T_{-+d}B_{-+} + f'_{-a}T_{++d}B_{+-} = f'_{+a}T_{++d}B_{--} + f'_{-a}T_{-+d}B_{++} \tag{5}$$

where T is the determinant of order two

$$T_d = \begin{vmatrix} f_d & g_d \\ f'_d & \Omega g'_d \end{vmatrix}$$

and where the first of the subscripts $\pm$ following T refers to the signs selecting the functions $f_\pm$ in the first column and the second subscript determines the functions $g_\pm$ in the second column. The quantities $B_{\pm\pm}$ are the expressions

$$B = f_b\psi_b H_1 f'_c - f_b\psi_b\psi_c H_2 f_c - f'_b H_3 f'_c + f'_b\psi_c H_4 f_c$$

and where, as before, the first of the $\pm$ subscripts determines the subscripts on the functions f_b and the second determines the subscripts on the functions f_c. As examples

$$T_{+-d} = \begin{vmatrix} f_{+d} & g_{-d} \\ f'_{+d} & \Omega g'_{-d} \end{vmatrix}$$

$$B_{-+} = f_{-b}\psi_b H_1 f'_{+c} - f_{-b}\psi_b\psi_c H_2 f_{+c} - f'_{-b} H_3 f'_{+c} + f'_{-b}\psi_c H_4 f_{+c}.$$

Since the region $b \leqslant z \leqslant c$ is assumed to be homogeneous

$$h_{\pm} = e^{\mp \zeta z}$$

where $\zeta^2 = (k^2 - \omega^2/v^2)$ and v is the shear wave velocity in this region. Thus

$$H_1 = -2\zeta \cosh \zeta(b-c) \qquad H_2 = 2\zeta^2 \sinh \zeta(b-c)$$
$$H_3 = -2 \sinh \zeta(b-c) \qquad H_4 = 2\zeta \cosh \zeta(b-c).$$

We now pass to the limit as $b \to c$. Let $\Delta = c-b$. Then

$$H_1 = -2\zeta(1+O(\Delta^2)) \qquad H_2 = -2\zeta^3\Delta(1+O(\Delta^2))$$
$$H_3 = 2\zeta\Delta(1+O(\Delta^2)) \qquad H_4 = 2\zeta(1+O(\Delta^2))$$
$$f_{+c} = f_{+b} + \Delta f'_{+b} + O(\Delta^2), \text{ etc.}$$
$$\psi_c = \psi_b + \Delta\psi'_b + O(\Delta^2).$$

Substituting into (5) and retaining terms to order Δ, we find

$$\tfrac{1}{2}B_{\pm\pm} = \zeta\psi(f'f - ff') + \Delta\zeta\{f'f'(\psi - 1) + \psi'f'f + ff\psi^2\zeta^2 - f''f\psi\} + O(\Delta^2) \qquad (6)$$

where again the $\pm$ subscript notation refers to the first and second of the function pairs ff respectively. As in our previous example

$$\tfrac{1}{2}B_{-+} = \zeta\psi(f'_-f_+ - f_-f'_+) + \Delta\zeta\{f'_-f'_+(\psi - 1) + \psi'f'_-f_+ + f_-f_+\psi^2\zeta^2 - f''_-f_+\psi\} + O(\Delta^2).$$

We write (6) in the form

$$\tfrac{1}{2}B_{\pm\pm} = \zeta\psi(f'f - ff') + \Delta\zeta\beta_{\pm\pm} + O(\Delta^2). \qquad (7)$$

We have dropped the subscript b from the expression. Substituting (7) into (5) we obtain, to terms of order Δ,

$$f'_{+a}T_{-+d} - f'_{-a}T_{++d} = \frac{\Delta[f'_{+a}(T_{++d}\beta_{--} - T_{-+d}\beta_{+}) + f'_{-a}(T_{-+d}\beta_{++} - T_{++d}\beta_{+-})]}{\psi(f'_-f_+ - f_-f'_+)}. \qquad (8)$$

We note that as $\Delta \to 0$ we have the usual eigenvalue equation (4) for Love waves. As a further check upon this result, we note that, as the properties of the layer of width Δ are made homogeneous with that of the original medium, we also obtain the usual eigenvalue equation (4) for Love waves. This last is demonstrated by observing that $\beta \to 0$ as $\psi \to 1$ and $\zeta \to \eta$ where we use the result of equation (2) and the condition of homogeneity of the layer in the proof.

Let it be assumed that (4) has a known solution for the simple unperturbed structure

$$L(k_0) = 0.$$

Then after the perturbation

$$\dot{L}(k_0)(k-k_0) = \frac{\Delta T_{++d}}{\psi}\frac{\left[\dfrac{(f'_{+a})^2\beta_{--} + (f'_{-a})^2\beta_{++}}{f'_{+a}} - f'_{-a}(\beta_{-+} + \beta_{+-})\right]}{f'_-f_+ - f_-f'_+} \qquad (9)$$

where the dot refers to differentiation with respect to k.

Because of the method of selecting the functions

$$f_+(z) = f_-(-z)$$

it follows that

$$-f'_+(z) = f'_-(-z)$$

where the primes still indicate differentiation with respect to $+z$. Thus equation (9) reduces to

$$\dot{L}(k_0)(k-k_0) = \frac{\Delta T_{++d}}{2\psi f'_- f_+}\left[\frac{(f'_+{}^a\beta_{--}+f'_-{}^a\beta_{++})}{f'_{+a}} - f'_{-a}(\beta_{-+}+\beta_{+-})\right] \qquad (10)$$

where it must be remembered that the expressions in parentheses are evaluated at $z = b$, and the functions f, T and B are evaluated for the solution k_0. When unspecified the functions are evaluated at $z = b$.

3. Perturbation theory: lower region

As before, we introduce a layer into the geometry; in this case the inhomogeneity is a homogeneous layer of thickness $\Delta = c-b$ located at $b \leqslant z \leqslant c$ where $b > d$. In the region $a \leqslant z \leqslant d$ we have solutions $f_\pm$; in the region $d \leqslant z \leqslant b$, we have solutions $g_\pm$; in the region below the layer $z \geqslant c$ the solution is $g_\pm$.

The eigenvalue problem is solved by the determination of the zeros of the equation

$$\begin{vmatrix} f'_{+a} & f'_{-a} & 0 & 0 & 0 & 0 & 0 \\ f_{+d} & f_{-d} & g_{+d} & g_{-d} & 0 & 0 & 0 \\ f'_{+d} & f'_{-d} & \Omega g'_{+d} & \Omega g'_{-d} & 0 & 0 & 0 \\ 0 & 0 & g_{+b} & g_{-b} & h_{+b} & h_{-b} & 0 \\ 0 & 0 & g'_{+b} & g'_{-b} & \psi_b h'_{+b} & \psi_b h'_{-b} & 0 \\ 0 & 0 & 0 & 0 & h_{+c} & h_{-c} & g_{+c} \\ 0 & 0 & 0 & 0 & \psi_c h'_{+c} & \psi_c h'_{-c} & g'_{+c} \end{vmatrix} = 0 \qquad (11)$$

We may expand (11) as before, obtaining

$$(f'_{+a}T_{-+d} - f'_{-a}T_{++d})A_{-+} = (-f'_{+a}T_{--d} + f'_{-a}T_{+-d})A_{++} \qquad (12)$$

where

$$A = g_b\psi_b H_1 g'_c - g_b\psi_b\psi_c H_2 g_c - g'_b H_3 g'_c + g'_b\psi_c H_4 g_c$$

with the functions H defined as before. Then, as before, to terms of order Δ,

$$(f'_{+a}T_{-+d} - f'_{-a}T_{++d}) = \frac{\Delta}{\psi}\frac{\alpha_{++}(f'_{-a}T_{+-d} - f'_{+a}T_{--d}) - \alpha_{-+}(f'_{+a}T_{-+d} - f'_{-a}T_{++d})}{g'_- g_+ - g_- g'_+} \qquad (13)$$

where

$$\alpha_{\pm\pm} = g'g(\psi-1) + \psi' g' g + gg\psi^2\zeta^2 - g''g\psi$$

with the sign convention as before.

By the same criterion as before

$$\dot{L}(k_0)(k-k_0) = \frac{\Delta\alpha_{++}}{2\psi g'_- g_+}(f'_- a T_{+-a} - f'_{+a} T_{--a}). \tag{14}$$

4. Green's function for Love waves

From equations (10) and (14) we can write the perturbation of the eigenvalue equation due to a continuous distribution of inhomogeneous properties $\psi(z)$ and $\zeta(z)/\eta(z)$. We obtain the integral expression

$$\begin{aligned} 2\dot{L}(k_0)(k-k_0) = {} & T_{++d} f'_{+a} \int_a^d \frac{\beta_{--}\,dz}{f'_- f_+ \psi} + \frac{T_{++d} f'_{-a}{}^2}{f'_{+a}} \int_a^d \frac{\beta_{++}\,dz}{f'_- f_+ \psi} \\ & - T_{++d} f'_{-a} \int_a^d \frac{(\beta_{-+}+\beta_{+-})\,dz}{f'_- f_+ \psi} + T_{+-a} f'_{-a} \int_d^\infty \frac{\alpha_{++}\,dz}{g'_- g_+ \psi} - T_{--a} f'_{+a} \int_d^\infty \frac{\alpha_{++}\,dz}{g'_- g_+ \psi} \end{aligned} \tag{15}$$

where the arguments of the integrals are functions of z. If the solution $f_\pm(z)$, $g_\pm(z)$ is known for a particular inhomogeneity, then a perturbation of this inhomogeneity $\psi(z)$ and $\zeta(z)/\eta(z)$ will yield the perturbation of the eigenvalues given by (15).

A typical term in the integrands, say β_{-+}/ψ, can be written as

$$\beta_{-+}/\psi = f'_- f_+(1-\psi^{-1}) + f'_- f_+(\psi\mu'/\mu + \psi'/\psi) + f_- f_+ \omega^2 \psi / v_0^2 - f_- f_+ \omega^2 \chi / v_0^2 \tag{16}$$

where v_0 is the unperturbed shear wave velocity $(\mu/\rho)^{\frac{1}{2}}$ at the appropriate depth and χ is the ratio of the density in the perturbed case to that in the unperturbed case. Hence the integrals can be separated into terms that depend only on the ratio of moduli ψ and into terms that depend only on the ratio of densities χ. The effects of variation of density and elastic modulus are thus decoupled and do not influence one another.

In order to avoid difficulties that may arise later, we linearize equation (16). We let the perturbation in density be $R(z)$ and the perturbation in elastic modulus be $M(z)$:

$$\begin{aligned} \chi &= 1 + R(z)/\rho(z), \qquad R/\rho \ll 1 \\ \psi &= 1 + M(z)/\mu(z), \qquad M/\mu \ll 1. \end{aligned}$$

Then

$$\beta_{-+}/\psi = f'_- f_+(M/\mu) + f'_- f_+(\mu'/\mu)(M/\mu) + f_- f_+(\omega^2/v_0^2)(M/\mu) - f_- f_+(\omega^2/v_0^2)(R/\rho). \tag{17}$$

We can thus compute the effect of the variation in density and the effect of the variation in modulus independently. These results are

$$\begin{aligned} 2\dot{L}(k_0)(k-k_0) = {} & -T_{++d} \int_a^d \frac{\omega^2}{v_0^2} \frac{R}{\rho} \frac{(f'_{+a} f_- - f'_{-a} f_+)^2}{f'_- f_+ f'_{+a}}\,dz - \\ & -(T_{+-a} f'_{-a} - T_{--a} f'_{+a}) \int_d^\infty \frac{g_+}{g'_-} \frac{\omega^2}{v_0^2} \frac{R}{\rho}\,dz \end{aligned} \tag{18}$$

for the case of variation in density, while for the variation in modulus it is

$$2\dot{L}(k_0)(k-k_0) = T_{+-a}\int_a^d \frac{M}{\mu}\times$$

$$\times\frac{\left\{\frac{\omega^2}{v_0^2}(f'_+af'_- - f'_-af'_+)^2+\frac{\mu'}{\mu}(f'_+af'_- - f'_-af'_+)(f'_+af'_- - f'_-af'_+)+(f'_+af'_- - f'_-af'_+)^2\right\}}{f'_-f'_+f'_+a}dz+ \tag{19}$$

$$+(T_{+-a}f'_{-a}-T_{--a}f'_{+a})\int_a^\infty \frac{M}{\mu}\frac{\left\{\frac{\omega^2}{v_0}g_+ + \frac{\mu'}{\mu}g'_+ + (g'_+)^2/g_+\right\}}{g'_-}dz.$$

5. First-order inversion of the perturbation expressions: density

The Green's function expressions (18) and (19) cannot, in general, be inverted analytically. Only for special solutions $f_\pm$ and $g_\pm$ can it be expected that the perturbations M/μ and R/ρ may be determined analytically for given dispersion data $\dot{L}(k_0)(k-k_0)$. Indeed, after a first inversion, it is not to be expected that new estimates of $\mu(z)$ and $\rho(z)$ will in turn lead to functions $f_\pm$ and $g_\pm$ that will be analytically describable. Thus, in general, the inversion of expressions (18) and (19) cannot be attempted except by lengthy and involved numerical procedures.

At least one case, however, is tractable. The familiar Love wave problem of a homogeneous layer overlying a homogeneous half-space leads to functions that permit a reduction of the expressions in question. Let

$$\begin{aligned} &\rho=\rho_1, \quad \mu=\mu_1 \qquad a<z<d \\ &\rho=\rho_2, \quad \mu=\mu_2 \qquad d<z<\infty \end{aligned} \qquad \Omega=\mu_2/\mu_1.$$

Then

$$f_\pm = \exp(\mp\eta_1 z), \qquad g_\pm = \exp(\mp\eta_2 z).$$

We further set $a=0$. Then equation (18) becomes

$$2\dot{L}(k_0)(k-k_0) = 4(\eta_1-\Omega\eta_2)\exp[-(\eta_1+\eta_2)d]\int_0^d \frac{\omega^2}{\mu_1}R\cosh^2\eta_1 z\,dz -$$

$$-\eta_1[(\Omega\eta_2+\eta_1)\exp[(\eta_2-\eta_1)d]+(\Omega\eta_2-\eta_1)\exp[(\eta_1+\eta_2)d]\int_d^\infty \frac{\exp(-2\eta_2 z)}{\eta_2}\frac{\omega^2}{\mu_2}R\,dz. \tag{20}$$

From equation (4)

$$L(k_0) = \eta_1(\Omega\eta_2+\eta_1)\exp[(\eta_1-\eta_2)d]+\eta_1(\Omega\eta_2-\eta_1)\exp[-(\eta_1+\eta_2)d] = 0. \tag{21}$$

Substituting this expression in equation (20) we obtain the equation

$$\left\{\Omega\left(\frac{\eta_2^2-\eta_1^2}{\Omega^2\eta_2^2-\eta_1^2}\right)+\eta_2 d\right\}k_0(k-k_0) = \eta_2\frac{\omega^2}{\mu_1}\int_0^d R\cosh^2\eta_1 z\,dz + \\ +\tfrac{1}{2}\eta_1\sinh 2\eta_1 d\left(\frac{\omega^2}{\mu_2}\right)\exp(2\eta_2 d)\int_d^\infty R\exp(-2\eta_2 z)\,dz. \tag{22}$$

Since η_1 is imaginary and η_2 is real, the above integral equation is real.

In order to invert this expression, it will be noted that

$$\int_0^\infty R\exp(-2\eta_2 z)dz$$

is the Laplace transform of $R(z)$ with the transform variable $2\eta_2$. Hence

$$R(z) = \left[\mathscr{L}^{-1}_{2\eta_2}\left\{\left(\frac{\Omega(\eta_2^2-\eta_1^2)}{\Omega^2\eta_2^2-\eta_1^2}+\eta_2 d\right)\frac{k_0(k-k_0)2\mu_2\exp(-2\eta_2 d)}{\omega^2\eta_1\sinh 2\eta_1 d}\right\}\right]H(z-d) \qquad z>d \tag{23}$$

is the required density perturbation in the half-space where the inverse Laplace transform variable is $2\eta_2$.

Similarly

$$\Phi(\eta_1) = \int_0^\infty R\cos^2(i\eta_1 z)\,dz = \tfrac{1}{2}\mathscr{C}_{2i\eta_1}\{R(z)\}+\tfrac{1}{2}\Phi(0)$$

where $\mathscr{C}_{i\eta_1}$ is the cosine transform with $2i\eta_1$ the transform variable. Thus

$$R(z) = \left[\mathscr{C}^{-1}_{2i\eta_1}\left\{\left(\frac{2\Omega(\eta_2^2-\eta_1^2)}{\Omega^2\eta_2^2-\eta_1^2}+\eta_2 d-\frac{1}{\Omega}\right)\frac{k_0(k-k_0)\mu_1}{\eta_2\omega^2}\right\}\right]H(d-z)H(z) \qquad 0<z<d \tag{24}$$

is the density perturbation in the overlying layer corresponding to the wave number perturbation $(k-k_0)$ where the inverse cosine transform variable is $2i\eta_1$, a real quantity.

6. First-order inversion of the perturbation expression: shear modulus

As a corollary to the preceding section it may be assumed that the wave number perturbation is entirely due to perturbation in the shear modulus. If the usual structure of a homogeneous layer overlying a homogeneous half-space is assumed, then equation (19) can be inverted. Equation (19) becomes

$$2\dot{L}(k_0)(k-k_0) = -4(\eta_1-\Omega\eta_2)\exp[-(\eta_1+\eta_2)d]\int_0^d \frac{M}{\mu_1}\left\{\frac{\omega^2\rho_1}{\mu_1}+k_0^2\sinh^2\eta_1 z\right\}dz + \\ +\eta_1\left[(\Omega\eta_2+\eta_1)\exp[(\eta_2-\eta_1)d]+(\Omega\eta_2-\eta_1)\exp[(\eta_2+\eta_1)d]\right]\int_d^\infty \frac{M}{\mu_2}\frac{k_0^2}{\eta_2}\exp(-2\eta_2 z)\,dz. \tag{25}$$

Applying the result of equation (21) we obtain the equation

$$\left(\Omega\frac{\eta_2^2-\eta_1^2}{\Omega^2\eta_2^2-\eta_1^2}+\eta_2 d\right)k_0(k-k_0) = -\int_0^d \eta_2\frac{M}{\mu_1}\left(\frac{\omega^2\rho_1}{\mu_1}+k_0^2\sinh^2\eta_1 z\right)dz - \tag{26}$$

$$-\tfrac{1}{2}\eta_1\sinh\eta_1 d\int_d^\infty \frac{M}{\mu_2}k_0^2\exp(-2\eta_2 z)\,dz.$$

As in the case of the density equation, this is a real integral equation.

The inversion proceeds in a manner similar to that presented above. We obtain, in the two regions,

$$M(z) = -\mu_2\mathscr{L}^{-1}_{2\eta_2}\left[\left(\frac{\Omega(\eta_2^2-\eta_1^2)}{\Omega^2\eta_2^2-\eta_1^2}+\eta_2 d\right)\frac{k-k_0}{k_0}\frac{2\exp(-2\eta_2 d)}{\eta_1\sinh\eta_1 d}\right]H(z-d) \tag{27}$$

$$z > d$$

$$M(z) = -\mu_1\mathscr{C}^{-1}_{2\eta_1}\left[\left(\frac{\Omega(\eta_2^2-\eta_1^2)}{\Omega^2\eta_2^2-\eta_1^2}-\frac{1}{\Omega}+\frac{k_0^2\mu_1}{2\omega^2\rho_1}\left(\frac{1}{\Omega}+\eta_2 d\right)\right)\frac{2(k-k_0)}{k_0\eta_2}\right]H(d-z)H(z) \tag{28}$$

$$0 < z < d$$

It will be noted that singularities in the Laplace transforms occur in equations (23) and (27) at $\eta_1 d = \frac{1}{2}n\pi$ (in η_2 units). In most cases the function $\exp(-2\eta_2 d)$ is very small at these points. Nevertheless the singularities produce exponential anomalies in the inversion at a sequence of depths corresponding to $n = 1, 2, 3 \ldots$ although these exponential anomalies are trivially small in practice. These singularities have been introduced by neglecting the interactions among the continuum of inhomogeneities. The wave functions on the right-hand sides of equations such as (17) and (18) should be the functions *after* the inhomogeneities have been introduced; they have been taken, as an approximation, to be the unperturbed wave functions.

7. First-order inversion: general solution

Since the inversion is not unique we investigate the nature of the coupling between the two independent solutions. Let $M_0(z)$ and $R_0(z)$ be the two "pure" solutions corresponding to the cases of zero perturbation by density and modulus respectively. These solutions are given in the preceding sections. We now write

$$k(\omega)-k_0(\omega) = [k(\omega)-k_0(\omega)+a(\omega)]-a(\omega)$$

where $a(\omega)$ is an arbitrary function, but one which has certain analytic properties. Let the term in brackets replace the term $(k-k_0)$ in equations (23) and (24) and $-a(\omega)$ replace $(k-k_0)$ in equations (27) and (28). From equation (17), the solution thus obtained is a solution to the problem of a perturbation $k-k_0$. We define the function

$$\phi(2\eta_2) = \left(\frac{\Omega(\eta_2^2-\eta_1^2)}{\Omega^2\eta_2^2-\eta_1^2}+\eta_2 d\right)\frac{2a(\omega)\exp(-2\eta_2 d)}{k_0\eta_1\sinh\eta_1 d}.$$

Then

$$R(z) = \left[R_0(z) + \mathscr{L}^{-1}_{2\eta_z}\left(\frac{\phi k_0^2}{\omega^2}\right)\right] H(z-d) \qquad z > d$$

and

$$M(z) = \mathscr{L}^{-1}_{2\eta_z}(\phi) H(z-d). \qquad z > d.$$

Hence the general values of $R(z)$ and $M(z)$ are connected through the linear integral equation

$$R(z) - R_0(z) = \int_d^z W(z-\xi) M(\xi)\, d\xi \qquad z > d \qquad (29)$$

where

$$W = \mathscr{L}^{-1}_{2\eta_z}(k_0^2/\omega^2).$$

In a similar way we can write the integral equation

$$M(z) - M_0(z) = \int_d^z V(z-\xi) R(\xi)\, d\xi \qquad (30)$$

where

$$V = \mathscr{L}^{-1}_{2\eta_z}(\omega^2/k_0^2).$$

The convolution solution for the upper layer is

$$R(z) - R_0(z) = \int_0^z W_1(z-\xi) M(\xi)\, d\xi \qquad z > d$$

where

$$W_1 = \mathscr{C}^{-1}_{2i\eta_1}\left\{\frac{k_0^2}{\omega^2} \frac{2\Omega(\eta_2^2 - \eta_1^2) + (\eta_2 d - \frac{1}{\Omega})(\Omega^2\eta_2^2 - \eta_1^2)}{2\Omega(\eta_2^2 - \eta_1^2) + \left\{-\frac{2}{\Omega} + \frac{k_0^2\mu_1}{\omega^2\rho_1}\left(\eta_2 d + \frac{1}{\Omega}\right)\right\}(\Omega^2\eta_2^2 - \eta_1^2)}\right\}. \qquad (31)$$

8. Numerical discussion

A number of numerical examples can be constructed as exercises in the application of the above calculations. These will be reserved for a later paper. Only one example will be given here.

We consider the two modulus perturbations:

$$M(z) = 0{\cdot}2\delta(z - 1{\cdot}0918d)$$

$$M(z) = \frac{1}{d}\{H(z-d) - H(z-1{\cdot}2d)\}. \qquad (32)$$

The integral

$$\int_d^\infty M(z)\exp(-2\eta_2 z)\,dz$$

can be computed in each case and the results compared. The functions to be compared are

$$0{\cdot}2\exp[-2\eta_2(1{\cdot}0918)d] \quad \text{and} \quad (2\eta_2 d)^{-1}[\exp(-2\eta_2 d)-\exp(-2{\cdot}4\eta_2 d)].$$

These functions differ by less than 1 per cent over the range $0 < 2\eta_2 d < 5$. Hence it is impossible, unless measurements of Love waves are made in the range 12 to 15 s to an accuracy of better than 1 per cent, to distinguish between the models of perturbation consisting of a uniform shear (or density) anomaly between 35 ($= d$) km and 42 km and a delta function anomaly at 38·21 km. Measurements at longer periods must be made more accurately than those at the short periods. Other anomalies are evidently indistinguishable from the two described above if they are made increasingly stronger and narrower than the step offset, each in turn clustering about the delta function. Successively greater accuracies are required for the measurements to distinguish between these possible models as the delta function is approached as a limit. Stronger anomalies, of course, perturb the dispersion curves more seriously than do weak anomalies. Both of the anomalies considered above, as possible end members of a suite of anomalies, produce dispersion curves that are perturbed from those corresponding to the initial assumption, but are within 1 per cent of one another.

9. Conclusions

Two conclusions may be drawn from the calculations presented in this paper. The direct inversion of surface dispersion curves into unique density—and modulus—depth distributions is impossible; a further relation specifying one or both of these is necessary. In the absence of such additional information one can only construct a suite of interrelated density and modulus perturbations, given by equations (29) and (30), all of which exactly yield the same dispersion curve. This set can be limited by certain physical arguments, but the lack of uniqueness in the solution remains a cardinal argument against the exact inversion of observed dispersion data.

If the data are admittedly inaccurate, then the number of permissible solutions is augmented by another parameter. This parameter describes the latitude that a permissible exact solution may be given in view of a known inaccuracy in the measurement.

In view of the present accuracy of observations of surface wave dispersion data, and in view of the non-uniqueness of the inversion of exact dispersion data, it may be concluded that the "fine structure" of the Earth's interior as determined from surface waves remains in doubt.

Institute of Geophysics,
University of California,
Los Angeles 24,
California.

1960 *May.*

References

Ewing, W. M., Jardetzky, W. S. & Press, F., 1957. *Elastic Waves in Layered Media* (McGraw-Hill, New York).

Jeffreys, H., 1924. On certain approximate solutions of linear differential equations of the second order, *Proc. Lond. Math. Soc.*, **23**, 428–436.

Jeffreys, H. & Jeffreys, B. S., 1956. *Methods of Mathematical Physics* (Cambridge, 3rd edition).

Morse, P. M. & Feshbach, H., 1953. *Methods of Theoretical Physics* (McGraw-Hill, New York).

Mott, N. F. & Massey, H. S. W., 1933. *The Theory of Atomic Collisions* (Oxford).

Takahashi, T., 1955. Analysis of the dispersion curves of Love waves, *Bull. Earthquake Research Inst.*, **33**, 287–296.

Takahashi, T., 1957. The dispersion of Rayleigh waves in heterogeneous media, *Bull. Earthquake Research Inst.*, **35**, 297–308.

The Transmission of a Rayleigh Pulse round a corner

E. R. Lapwood

"At the round world's imagined corners."—*John Donne, Holy Sonnets, No. 7*

Summary

Lamb, in his classical paper of 1904, found the form of the Rayleigh pulse that travels over the surface of an elastic half-space, which has been struck along a line. Let such a pulse travel on one face of an elastic quarter-space, the crest-line being parallel to and travelling towards the edge. We enquire what form the Rayleigh pulse will have after passing round the corner.

The problem is taken as two-dimensional; the elastic solid fills the positive quadrant xOy and the incident Rayleigh pulse arises from a pressure

$$p(x,t) = -Q\delta(x-a)\Phi(t)$$

applied on the face $y = 0$.

The differential equations and boundary conditions are reduced to operational form by the use of a Fourier transform with respect to time and Laplace transforms with respect to x and y. Difficulties arise from our lack of knowledge of the displacements at the surfaces $x = 0$ and $y = 0$, but solution of the problem is reduced to that of two simultaneous integral equations.

An iterative solution can be found, and from this we isolate the parts which describe the incident Rayleigh pulse (identical with Lamb's solution) and the pulse transmitted round the corner.

It is found that the form of the pulse is greatly changed, in ways depending on the values of the velocities of P, S and Rayleigh waves. Whereas the dispacements u and v on $y = 0$ have the shape of $\Phi(t)$ and its *allied function* (Hilbert transform) $\Phi'(t)$ respectively, each of u and v on $x = 0$ is given by a linear combination of Φ and Φ'. Since Φ' may differ greatly from Φ in form, the change in shape of each component of displacement when it turns the corner may be very marked.

Introduction

Observations of the advance of Rayleigh waves across continents and oceans lead naturally to the question in mathematical theory: "What happens when an advancing Rayleigh wave impinges on a different medium?" Reducing the problem to two dimensions, we think of two quarter-spaces of different materials,

welded together, and ask how a Rayleigh wave is affected as it passes from one to the other. This is a boundary value problem with eight boundary conditions. Unfortunately, when the advancing Rayleigh wave is expressed in terms of potentials of dilatational and distortional displacement, the two potentials vary exponentially with depth, and at different rates. It proves to be difficult to guess a set of potential functions suitable for the satisfying of the boundary conditions. Indeed, further consideration shows that it is unlikely that this is the right mode of attack. For it is very unlikely that the system—advancing Rayleigh wave, transmitted Rayleigh wave, reflected Rayleigh wave—will describe the phenomenon completely. May not an interface wave travel down the common boundary of the two media, and will not both dilatational and distortional body waves be radiated from the neighbourhood of the corner?

The likelihood of such effects makes it necessary to seek a different specification of the problem. Dr R. Stoneley suggested (in a private communication) that it would be better to consider the complete field of disturbance after the emission of a pulse from a source in one quarter-space. Then, in an attack by classical methods, Lamb's problem (1904) and its solutions offer a natural starting point and check. But we still find ourselves faced with the need to construct potential functions to satisfy the boundary conditions on the interface. Whereas in Lamb's problem for a buried source there is a single reflection only and potentials can be guessed, here, if we regard the problem as one in geometrical ray theory, we have reflections or refractions at three boundaries, and if we apply Huyghens' principle we have to admit an infinite system of reflected and refracted rays.

The problem is first simplified, therefore, to become that of transmission of a Rayleigh pulse around the corner of a quarter-space. An attack by the construction of the successive reflections of an outgoing wave still runs into great complexity of multiple integrals when the second boundary is reached, and so the problem is worked by operational methods. Such methods were first applied to problems of seismology by Sir Harold Jeffreys (1926, 1931), and I am very glad to be able to offer this contribution to the volume published in his honour.

1. The problem set up

Let a homogeneous isotropic quarter-space $x \geqslant 0$, $y \geqslant 0$ be struck along the line $x = a$, $y = 0$. We want to find the subsequent disturbance throughout the quarter-space, and in particular to discuss the Rayleigh pulses on Ox and Oy.

In a standard notation, the equations of motion are

$$\left.\begin{aligned} \rho u_{tt} &= (p_{xx})_x + (p_{xy})_y, \\ \rho v_{tt} &= (p_{xy})_x + (p_{yy})_y, \end{aligned}\right\} \tag{1.1}$$

where (u, v) is the displacement at (x, y) and suffixes, except for those in the stress tensor, indicate partial derivatives. The stress–strain relations are

$$\left.\begin{aligned} p_{xx} &= (\lambda + 2\mu)u_x + \lambda v_y, \\ p_{yy} &= \lambda u_x + (\lambda + 2\mu)v_y, \\ p_{xy} &= \mu(u_y + v_x). \end{aligned}\right\} \tag{1.2}$$

We take first a more general boundary condition, namely that on $y = 0$ the normal stress is given by a known function $p(x, t)$. All other boundary stresses are

zero. Thus we have

$$\left.\begin{aligned} p_{xx} &= 0, \quad p_{xy} = 0 \qquad \text{on } x = 0 \text{ for all } t, \\ p_{xy} &= 0, \quad p_{yy} = p(x,t) \quad \text{on } y = 0 \text{ for all } t. \end{aligned}\right\} \tag{1.3}$$

We replace the Lamé constants λ and μ by the dilatational and shear wave velocities, writing

$$\lambda + 2\mu = \rho\alpha^2, \qquad \mu = \rho\beta^2.$$

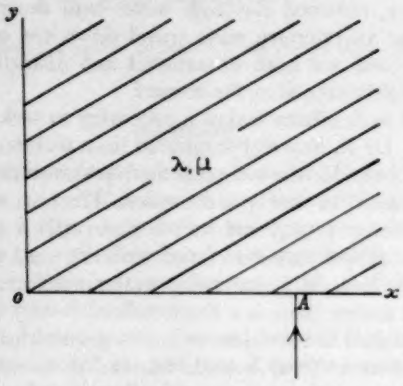


FIG. 1.—The elastic quarter-space.

Then using (1.2) and (1.1) we obtain

$$\left.\begin{aligned} u_{tt} &= \alpha^2 u_{xx} + (\alpha^2 - \beta^2) v_{xy} + \beta^2 u_{yy}, \\ v_{tt} &= \beta^2 v_{xx} + (\alpha^2 - \beta^2) u_{xy} + \alpha^2 v_{yy}, \end{aligned}\right\} \tag{1.4}$$

with

$$\left.\begin{aligned} \alpha^2 u_x + (\alpha^2 - 2\beta^2) v_y &= 0, \\ u_y + v_x &= 0, \end{aligned}\right\} \text{ on } x = 0 \text{ for all } t,$$
$$\left.\begin{aligned} u_y + v_x &= 0, \\ (\alpha^2 - 2\beta^2) u_x + \alpha^2 v_y &= p(x,t)/\rho, \end{aligned}\right\} \text{ on } y = 0 \text{ for all } t. \tag{1.5}$$

We note that the form of u and v will be expected to differ on different parts of $y = 0$. In particular, a pulse from a blow at $x = a$ will be travelling in opposite directions in $x < a$ and $x > a$.

2. Transformation of the equations into operational form

Let f^ω denote the Fourier Transform of $f(t)$, i.e.

$$f^\omega = \mathcal{F}\{f(t)\} = \frac{1}{\sqrt{(2\pi)}} \int_{-\infty}^{\infty} f(t) \exp(i\omega t)\, dt. \tag{2.1}$$

Then, provided $f(t) \to 0$ and $\partial f/\partial t \to 0$ as $t \to \pm\infty$,

$$\mathcal{F}(f_t) = -i\omega f^\omega, \qquad \mathcal{F}(f_{tt}) = -\omega^2 f^\omega.$$

Assuming that u and v have these properties, we transform (1.4) and (1.5) with regard to t and obtain

$$\left.\begin{aligned} \alpha^2 u_{xx}^{\omega} + (\alpha^2-\beta^2) v_{xy}^{\omega} + \beta^2 u_{yy}^{\omega} &= -\omega^2 u^{\omega}, \\ \beta^2 v_{xx}^{\omega} + (\alpha^2-\beta^2) u_{xy}^{\omega} + \alpha^2 v_{yy}^{\omega} &= -\omega^2 v^{\omega}, \end{aligned}\right\} \tag{2.2}$$

$$\left.\begin{aligned} \alpha^2 u_x^{\omega} + (\alpha^2-2\beta^2) v_y^{\omega} &= 0, \\ u_y^{\omega} + v_x^{\omega} &= 0, \end{aligned}\right\} \text{ on } x = 0, \qquad \left.\begin{aligned} u_y^{\omega} + v_x^{\omega} &= 0, \\ (\alpha^2-2\beta^2) u_x^{\omega} + \alpha^2 v_y^{\omega} &= p^{\omega}(x)/\rho, \end{aligned}\right\} \text{ on } y = 0. \tag{2.3}$$

Next let f^{ξ} denote the Laplace Transform of $f(x)$, i.e.

$$f^{\xi} = \mathscr{L}\{f(x)\} = \int_0^{\infty} f(x) \exp(-\xi x)\, dx.$$

Then, with the usual assumptions about the behaviour of $f(x)$ as $x \to \infty$,

$$\left.\begin{aligned} \mathscr{L}\{f_x\} &= \xi f^{\xi} - f(0), \\ \mathscr{L}\{f_{xx}\} &= \xi^2 f^{\xi} - \xi f(0) - f_x(0). \end{aligned}\right\} \tag{2.4}$$

We assume that u and v obey these assumptions. In transforming (2.2) and (2.3) with regard to x we must retain the unknown functions

$$u^{\omega}(0, y, \omega), \quad u_x^{\omega}(0, y, \omega), \quad v^{\omega}(0, y, \omega), \quad v_x^{\omega}(0, y, \omega),$$

and we have

$$\left.\begin{aligned} &\alpha^2\{\xi^2 u^{\omega\xi} - \xi u^{\omega}(0, y, \omega) - u_x^{\omega}(0, y, \omega)\} + \\ &\quad + (\alpha^2-\beta^2)\{\xi v_y^{\omega\xi}(\xi, y, \omega) - v_y^{\omega}(0, y, \omega)\} + \beta^2 u_{yy}^{\omega\xi} = -\omega^2 u^{\omega\xi}, \\ &\beta^2\{\xi^2 v^{\omega\xi} - \xi v^{\omega}(0, y, \omega) - v_x^{\omega}(0, y, \omega)\} + \\ &\quad + (\alpha^2-\beta^2)\{\xi u_y^{\omega\xi}(\xi, y, \omega) - u_y^{\omega}(0, y, \omega)\} + \alpha^2 v_{yy}^{\omega\xi} = -\omega^2 v^{\omega\xi}, \end{aligned}\right\} \tag{2.5}$$

with

$$\left.\begin{aligned} \alpha^2 u_x^{\omega}(0, y, \omega) + (\alpha^2-2\beta^2) v_y^{\omega}(0, y, \omega) &= 0, \\ u_y^{\omega}(0, y, \omega) + v_x^{\omega}(0, y, \omega) &= 0, \end{aligned}\right\} \tag{2.6a}$$

$$\left.\begin{aligned} u_y^{\omega\xi}(\xi, 0, \omega) + \xi v^{\omega\xi}(\xi, 0, \omega) - v^{\omega}(0, 0, \omega) &= 0, \\ (\alpha^2-2\beta^2)\{\xi u^{\omega\xi}(\xi, 0, \omega) - u(0, 0, \omega)\} + \alpha^2 v_y^{\omega\xi}(\xi, 0, \omega) &= p^{\omega\xi}/\rho. \end{aligned}\right\} \tag{2.6b}$$

(2.6a) were not transformed with regard to x, since they refer to the boundary $x = 0$.

We finally transform with respect to y, using the notation

$$f^{\eta} = \mathscr{L}\{f(y)\} = \int_0^{\infty} f(y) \exp(-\eta y)\, dy, \tag{2.7}$$

and making the necessary assumptions about the behaviour of u and v as functions of y. We note that (2.6b) stand unchanged, and obtain

$$\begin{aligned} &\alpha^2\{\xi^2 u^{\omega\xi\eta} - \xi u^{\omega\eta}(0, \eta, \omega) - u_x^{\omega\eta}(0, \eta, \omega)\} + \\ &\quad + (\alpha^2-\beta^2)[\xi\{\eta v^{\omega\xi\eta}(\xi, \eta, \omega) - v^{\omega\xi}(\xi, 0, \omega)\} - \{\eta v^{\omega\eta}(0, \eta, \omega) - v^{\omega}(0, 0, \omega)\}] + \\ &\quad + \beta^2\{\eta^2 u^{\omega\xi\eta} - \eta u^{\omega\xi}(\xi, 0, \omega) - u_y^{\omega\xi}(\xi, 0, \omega)\} = -\omega^2 u^{\omega\xi\eta}, \end{aligned} \tag{2.8a}$$

$$\beta^2\{\xi^2 v^{\omega\xi\eta} - \xi v^{\omega\eta}(0,\eta,\omega) - v_x{}^{\omega\eta}(0,\eta,\omega)\} +$$
$$+ (\alpha^2-\beta^2)[\xi\{\eta u^{\omega\xi\eta}(\xi,\eta,\omega) - u^{\omega\xi}(\xi,0,\omega)\} - \{\eta u^{\omega\eta}(0,\eta,\omega) - u^{\omega}(0,0,\omega)\}] +$$
$$+ \alpha^2\{\eta^2 v^{\omega\xi\eta} - \eta v^{\omega\xi}(\xi,0,\omega) - v_y{}^{\omega\xi}(\xi,0,\omega)\} = -\omega^2 v^{\omega\xi\eta}, \tag{2.8b}$$

with

$$\left.\begin{aligned} \alpha^2 u_x{}^{\omega\eta}(0,\eta,\omega) + (\alpha^2-2\beta^2)\{\eta v^{\omega\eta}(0,\eta,\omega) - v^{\omega}(0,0,\omega)\} &= 0 \\ \eta u^{\omega\eta}(0,\eta,\omega) - u^{\omega}(0,0,\omega) + v_x{}^{\omega\eta}(0,\eta,\omega) &= 0, \end{aligned}\right\} \tag{2.9a}$$

$$\left.\begin{aligned} u_y{}^{\omega\xi}(\xi,0,\omega) + \xi v^{\omega\xi}(\xi,0,\omega) - v^{\omega}(0,0,\omega) &= 0, \\ (\alpha^2-2\beta^2)\{\xi u^{\omega\xi}(\xi,0,\omega) - u(0,0,\omega)\} + \alpha^2 v_y{}^{\omega\xi}(\xi,0,\omega) &= p^{\omega\xi}/\rho. \end{aligned}\right\} \tag{2.9b}$$

From (2.9), writing u_ξ, v_ξ for $u^{\omega\xi}(\xi,0,\omega)$, $v^{\omega\xi}(\xi,0,\omega)$, u_η, v_η for $u^{\omega\eta}(0,\eta,\omega)$, $v^{\omega\eta}(0,\eta,\omega)$, and u_0, v_0, for $u^{\omega}(0,0,\omega)$, $v^{\omega}(0,0,\omega)$, we get

$$\left.\begin{aligned} \alpha^2 u_x{}^{\omega\eta}(0,\eta,\omega) &= -(\alpha^2-2\beta^2)(\eta v_\eta - v_0), \\ v_x{}^{\omega\eta}(0,\eta,\omega) &= -(\eta u_\eta - u_0), \\ u_y{}^{\omega\xi}(\xi,0,\omega) &= -(\xi v_\xi - v_0), \\ \alpha^2 v_y{}^{\omega\xi}(\xi,0,\omega) &= -(\alpha^2-2\beta^2)(\xi u_\xi - u_0) + p^{\omega\xi}/\rho. \end{aligned}\right\} \tag{2.10}$$

We substitute these into (2.8) and collect terms to get

$$\left.\begin{aligned} (\alpha^2\xi^2+\beta^2\eta^2+\omega^2)u^{\omega\xi\eta} + (\alpha^2-\beta^2)\xi\eta v^{\omega\xi\eta} \\ = \beta^2\eta u_\xi + \alpha^2\xi u_\eta + (\alpha^2-2\beta^2)\xi v_\xi + \beta^2\eta v_\eta, \\ (\alpha^2-\beta^2)\xi\eta u^{\omega\xi\eta} + (\beta^2\xi^2+\alpha^2\eta^2+\omega^2)v^{\omega\xi\eta} \\ = \beta^2\xi u_\xi + (\alpha^2-2\beta^2)\eta u_\eta + \alpha^2\eta v_\xi + \beta^2\xi v_\eta + p^{\omega\xi}/\rho. \end{aligned}\right\} \tag{2.11}$$

We solve (2.11) for $u^{\omega\xi\eta}$, $v^{\omega\xi\eta}$ and find

$$\left.\begin{aligned} \Delta u^{\omega\xi\eta} &= (\beta^2\xi^2+\alpha^2\eta^2+\omega^2)\{\beta^2\eta u_\xi + \alpha^2\xi u_\eta + (\alpha^2-2\beta^2)\xi v_\xi + \beta^2\eta v_\eta\} - \\ &\quad -(\alpha^2-\beta^2)\xi\eta\{\beta^2\xi u_\xi + (\alpha^2-2\beta^2)\eta u_\eta + \alpha^2\eta v_\xi + \beta^2\xi v_\eta + p^{\omega\xi}/\rho\}, \\ \Delta v^{\omega\xi\eta} &= (\alpha^2\xi^2+\beta^2\eta^2+\omega^2)\{\beta^2\xi u_\xi + (\alpha^2-2\beta^2)\eta u_\eta + \alpha^2\eta v_\xi + \beta^2\xi v_\eta + p^{\omega\xi}/\rho\} - \\ &\quad -(\alpha^2-\beta^2)\xi\eta\{\beta^2\eta u_\xi + \alpha^2\xi u_\eta + (\alpha^2-2\beta^2)\xi v_\xi + \beta^2\eta v_\eta\}, \end{aligned}\right\} \tag{2.12}$$

where

$$\begin{aligned} \Delta &= (\alpha^2\xi^2+\beta^2\eta^2+\omega^2)(\beta^2\xi^2+\alpha^2\eta^2+\omega^2) - (\alpha^2-\beta^2)^2\xi^2\eta^2 \\ &= (\alpha^2\xi^2+\alpha^2\eta^2+\omega^2)(\beta^2\xi^2+\beta^2\eta^2+\omega^2). \end{aligned} \tag{2.13}$$

The equations (2.12) for $u^{\omega\xi\eta}$ and $v^{\omega\xi\eta}$ may be re-written:

$$\begin{aligned} u^{\omega\xi\eta} = u_\xi\eta\left\{\frac{\beta^2}{\beta^2(\xi^2+\eta^2)+\omega^2} - \frac{2\beta^2(\alpha^2-\beta^2)\xi^2}{\Delta}\right\} + \\ + u_\eta\xi\left\{\frac{\alpha^2}{\alpha^2(\xi^2+\eta^2)+\omega^2} + \frac{2\beta^2(\alpha^2-\beta^2)\eta^2}{\Delta}\right\} + \\ + v_\xi\xi\left\{\frac{\alpha^2-2\beta^2}{\alpha^2(\xi^2+\eta^2)+\omega^2} - \frac{2\beta^2(\alpha^2-\beta^2)\eta^2}{\Delta}\right\} + \\ + v_\eta\eta\left\{\frac{\beta^2}{\beta^2(\xi^2+\eta^2)+\omega^2} - \frac{2\beta^2(\alpha^2-\beta^2)\xi^2}{\Delta}\right\} - \\ - \frac{(\alpha^2-\beta^2)\xi\eta}{\rho\Delta}p^{\omega\xi}, \end{aligned} \tag{2.14}$$

$$v^{\omega\xi\eta} = u_\xi\xi\left\{\frac{\beta^2}{\beta^2(\xi^2+\eta^2)+\omega^2} - \frac{2\beta^2(\alpha^2-\beta^2)\eta^2}{\Delta}\right\} +$$
$$+ u_\eta\eta\left\{\frac{\alpha^2-2\beta^2}{\alpha^2(\xi^2+\eta^2)+\omega^2} - \frac{2\beta^2(\alpha^2-\beta^2)\xi^2}{\Delta}\right\} +$$
$$+ v_\xi\eta\left\{\frac{\alpha^2}{\alpha^2(\xi^2+\eta^2)+\omega^2} + \frac{2\beta^2(\alpha^2-\beta^2)\xi^2}{\Delta}\right\} +$$
$$+ v_\eta\xi\left\{\frac{\beta^2}{\beta^2(\xi^2+\eta^2)+\omega^2} - \frac{2\beta^2(\alpha^2-\beta^2)\eta^2}{\Delta}\right\} +$$
$$+ \frac{p^{\omega\xi}}{\rho}\left\{\frac{1}{\beta^2(\xi^2+\eta^2)+\omega^2} - \frac{(\alpha^2-\beta^2)\eta^2}{\Delta}\right\}. \tag{2.15}$$

3. Boundary conditions at $y = \infty$

The solution of the problem must be bounded as $y \to \infty$ and as $x \to \infty$. To consider the first of these requirements we must invert the expressions (2.14) and (2.15) with respect to η; y is then recovered, and we can discuss $u^{\omega\xi}$ and $v^{\omega\xi}$ as functions of y.

We now introduce the auxiliary parameters

$$\kappa_\alpha = \omega/\alpha, \qquad \kappa_\beta = \omega/\beta, \tag{3.1}$$

and write

$$\xi^2+\kappa_\alpha{}^2 = \mu_\alpha{}^2, \qquad \xi^2+\kappa_\beta{}^2 = \mu_\beta{}^2. \tag{3.2}$$

The definitions of μ_α, μ_β will be completed later. Then (2.14) and (2.15) become

$$u^{\omega\xi\eta} = u_\xi\left\{\frac{\eta}{\eta^2+\mu_\beta{}^2} - \frac{2(1-\beta^2/\alpha^2)\xi^2\eta}{\Delta_\eta}\right\} +$$
$$+ u_\eta\left\{\frac{\xi}{\eta^2+\mu_\alpha{}^2} + \frac{2(1-\beta^2/\alpha^2)\xi\eta^2}{\Delta_\eta}\right\} +$$
$$+ v_\xi\left\{\frac{(1-2\beta^2/\alpha^2)\xi}{\eta^2+\mu_\alpha{}^2} - \frac{2(1-\beta^2/\alpha^2)\xi\eta^2}{\Delta_\eta}\right\} +$$
$$+ v_\eta\left\{\frac{\eta}{\eta^2+\mu_\beta{}^2} - \frac{2(1-\beta^2/\alpha^2)\xi^2\eta}{\Delta_\eta}\right\} -$$
$$- \frac{(1-\beta^2/\alpha^2)\xi\eta}{\rho\beta^2\Delta_\eta}p^{\omega\xi}, \tag{3.3}$$

$$v^{\omega\xi\eta} = u_\xi\left\{\frac{\xi}{\eta^2+\mu_\beta{}^2} - \frac{2(1-\beta^2/\alpha^2)\xi\eta^2}{\Delta_\eta}\right\} +$$
$$+ u_\eta\left\{\frac{(1-2\beta^2/\alpha^2)\eta}{\eta^2+\mu_\alpha{}^2} - \frac{2(1-\beta^2/\alpha^2)\xi^2\eta}{\Delta_\eta}\right\} +$$
$$+ v_\xi\left\{\frac{\eta}{\eta^2+\mu_\alpha{}^2} + \frac{2(1-\beta^2/\alpha^2)\xi^2\eta}{\Delta_\eta}\right\} +$$
$$+ v_\eta\left\{\frac{\xi}{\eta^2+\mu_\beta{}^2} - \frac{2(1-\beta^2/\alpha^2)\xi\eta^2}{\Delta_\eta}\right\} +$$
$$+ \frac{p^{\omega\xi}}{\rho\beta^2}\left\{\frac{1}{\eta^2+\mu_\beta{}^2} - \frac{(1-\beta^2/\alpha^2)\eta^2}{\Delta_\eta}\right\}, \tag{3.4}$$

where

$$\Delta_\eta = (\eta^2+\mu_\alpha^2)(\eta^2+\mu_\beta^2). \tag{3.5}$$

Inverting the Laplace transforms $u^{\omega\xi\eta}$ and $v^{\omega\xi\eta}$ we recall that u_ξ and v_ξ are not functions of η, and note that

$$\frac{\eta}{\Delta_\eta} = \left(\frac{\eta}{\eta^2+\mu_\alpha^2} - \frac{\eta}{\eta^2+\mu_\beta^2}\right)\frac{1}{\kappa_\beta^2-\kappa_\alpha^2}, \tag{3.6}$$

$$\frac{\eta^2}{\Delta_\eta} = \left(-\frac{\mu_\alpha^2}{\eta^2+\mu_\alpha^2} + \frac{\mu_\beta^2}{\eta^2+\mu_\beta^2}\right)\frac{1}{\kappa_\beta^2-\kappa_\alpha^2}, \tag{3.7}$$

and

$$\mathscr{L}^{-1}\left(\frac{\eta}{\eta^2+\mu_\alpha^2}\right) = \cos\mu_\alpha y, \qquad \mathscr{L}^{-1}\left(\frac{\mu_\alpha}{\eta^2+\mu_\alpha^2}\right) = \sin\mu_\alpha y,$$

$$\mathscr{L}^{-1}\left\{f^\eta(\eta)\frac{\eta}{\eta^2+\mu_\alpha^2}\right\} = \int_0^y f(\nu)\cos\mu_\alpha(y-\nu)\,d\nu.$$

Then

$$u^{\omega\xi}(\xi, y, \omega) = u_\xi\cos\mu_\beta y - \frac{2\xi^2 u_\xi}{\kappa_\beta^2}(\cos\mu_\alpha y - \cos\mu_\beta y) +$$

$$+\frac{\xi}{\mu_\alpha}\int_0^y u^\omega(0,\nu,\omega)\sin\mu_\alpha(y-\nu)\,d\nu -$$

$$-\frac{2\xi}{\kappa_\beta^2}\int_0^y u^\omega(0,\nu,\omega)\{\mu_\alpha\sin\mu_\alpha(y-\nu) - \mu_\beta\sin\mu_\beta(y-\nu)\}\,d\nu +$$

$$+v_\xi(1-2\beta^2/\alpha^2)\frac{\xi}{\mu_\alpha}\sin\mu_\alpha y + \frac{2\xi v_\xi}{\kappa_\beta^2}(\mu_\alpha\sin\mu_\alpha y - \mu_\beta\sin\mu_\beta y) +$$

$$+\int_0^y v^\omega(0,\nu,\omega)\cos\mu_\beta(y-\nu)\,d\nu -$$

$$-\frac{2\xi^2}{\kappa_\beta^2}\int_0^y v^\omega(0,\nu,\omega)\{\cos\mu_\alpha(y-\nu) - \cos\mu_\beta(y-\nu)\}\,d\nu -$$

$$-\frac{\xi}{\rho\omega^2}p^{\omega\xi}(\cos\mu_\alpha y - \cos\mu_\beta y). \tag{3.8}$$

We want to ensure that this expression for $u^{\omega\xi}(\xi, y, \omega)$ is bounded as $y \to \infty$. We note that since ξ, when regarded as the variable in the inversion formula, lies on a Bromwich contour, and ω is taken to be real in the Fourier inversion formula, μ_α^2 and μ_β^2 are complex quantities except when $\mathscr{I}(\eta) = 0$. We first consider

the case $\mathscr{I}(\mu_\alpha) > 0$, $\mathscr{I}(\mu_\beta) > 0$. Then as $y \to \infty$,

$$\begin{aligned} \exp(i\mu_\alpha y) &\to 0, \qquad \exp(-i\mu_\alpha y) \to \infty, \\ \exp(i\mu_\beta y) &\to 0, \qquad \exp(-i\mu_\beta y) \to \infty, \end{aligned} \tag{3.9}$$

and for $u^{\omega\xi}(\xi, y, \omega)$ to be bounded as $y \to \infty$ it is necessary that the coefficients of $\exp(-i\mu_\alpha y)$ and $\exp(-i\mu_\beta y)$ in (3.8) vanish as $y \to \infty$. Thus

$$\int_0^\infty \exp(i\mu_\alpha \nu)\{(\kappa_\beta^2 - 2\kappa_\alpha^2 - 2\xi^2)u^\omega(0, \nu, \omega) + 2i\xi\mu_\alpha v^\omega(0, \nu, \omega)\}d\nu +$$

$$+ 2i\xi\mu_\alpha u_\xi + (\kappa_\beta^2 + 2\xi^2)v_\xi = -\frac{i\mu_\alpha}{\rho\beta^2}p^{\omega\xi}, \tag{3.10}$$

and

$$\int_0^\infty \exp(i\mu_\beta \nu)\{2i\xi\mu_\beta u^\omega(0, \nu, \omega) + (\kappa_\beta^2 + 2\xi^2)v^\omega(0, \nu, \omega)\}d\nu +$$

$$+ (\kappa_\beta^2 + 2\xi^2)u_\xi - 2i\xi\mu_\beta v_\xi = -\frac{\xi}{\rho\beta^2}p^{\omega\xi}. \tag{3.11}$$

If we follow through the same argument applied to $v^{\omega\xi}(\xi, y, \omega)$ in place of $u^{\omega\xi}(\xi, y, \omega)$ we again obtain the equations (3.10) and (3.11).

Regarding (3.10) and (3.11) as equations for u_ξ and v_ξ we may solve to obtain, for $\mathscr{I}(\mu_\alpha) > 0, \mathscr{I}(\mu_\beta) > 0$.

$$F(\xi)u^{\omega\xi}(\xi, 0, \omega) = (-\kappa_\beta^2 - 2\xi^2 + 2\mu_\alpha\mu_\beta)\frac{\xi p^{\omega\xi}}{\rho\beta^2} - 2i\xi\mu_\beta I_\alpha - (\kappa_\beta^2 + 2\xi^2)I_\beta, \tag{3.12}$$

$$F(\xi)v^{\omega\xi}(\xi, 0, \omega) = -\frac{i\mu_\alpha\kappa_\beta^2}{\rho\beta^2}p^{\omega\xi} - (\kappa_\beta^2 + 2\xi^2)I_\alpha + 2i\xi\mu_\alpha I_\beta, \tag{3.13}$$

where

$$F(\xi) = (\kappa_\beta^2 + 2\xi^2)^2 - 4\xi^2\mu_\alpha\mu_\beta, \tag{3.14}$$

$$I_\alpha = \int_0^\infty \exp(i\mu_\alpha \nu)\{(\kappa_\beta^2 - 2\kappa_\alpha^2 - 2\xi^2)u^\omega(0, \nu, \omega) + 2i\xi\mu_\alpha v^\omega(0, \nu, \omega)\}d\nu, \tag{3.15}$$

and

$$I_\beta = \int_0^\infty \exp(i\mu_\beta \nu)\{2i\xi\mu_\beta u^\omega(0, \nu, \omega) + (\kappa_\beta^2 + 2\xi^2)v^\omega(0, \nu, \omega)\}d\nu. \tag{3.16}$$

It is a condition on the use of the inversion formula in the operational method that there should be no singularity of the integrand on the right of the Bromwich contour. Thus any cuts which are made in the ξ plane in order to make μ_α, μ_β single valued must allow for a Bromwich contour which does not meet them. This rules out the Sommerfeld method in which the top sheet of the Riemann surface

is defined by $\mathscr{I}(\mu_\alpha) > 0$, $\mathscr{I}(\mu_\beta) > 0$, since there the cuts (for real ω) extend all along the real axis of ξ. Instead, we cut the ξ-plane along straight lines from the branch points $\pm i\kappa_\alpha$, $\pm i\kappa_\beta$ running parallel to the negative real axis to $-\infty$. These will not interfere with any contour $(c-i\infty, c+i\infty)$, where $c > 0$, but they will make it necessary for us to define μ_α and μ_β appropriately. We therefore take μ_α, μ_β to be positive when ξ lies on the positive real axis. This means that when ξ moves from the real axis up the Bromwich contour in the first quadrant $\mathscr{I}(\mu_\alpha)$, $\mathscr{I}(\mu_\beta)$ are positive, but when ξ moves down the Bromwich contour in the fourth quadrant $\mathscr{I}(\mu_\alpha)$, $\mathscr{I}(\mu_\beta)$ are negative. We must therefore consider the values taken by $u^{\omega\xi}(\xi, y, \omega)$, $v^{\omega\xi}(\xi, y, \omega)$ when $\mathscr{I}(\mu_\alpha) < 0$, $\mathscr{I}(\mu_\beta) < 0$. We find that we obtain again the conditions (3.10) and (3.11) with the exception that $+i\mu_\alpha$ and $+i\mu_\beta$ are replaced by $-i\mu_\alpha$ and $-i\mu_\beta$ at each occurrence. Then our final equations (3.12) to (3.16) remain true provided the same replacements are made.

It would appear from this that there are sudden changes of sign, and therefore discontinuities, in the expressions for $u^{\omega\xi}(\xi, 0, \omega)$ and $v^{\omega\xi}(\xi, 0, \omega)$ as ξ crosses the real axis. But the terms in (3.10) and (3.11) have each been obtained by taking one exponential term out of the pair in a sine or cosine. Which of the pair is chosen depends on which dominates the expression for large y; different terms will be taken according as $\mathscr{I}(\mu_\alpha)$ and $\mathscr{I}(\mu_\beta)$ are greater than or less than zero. As $\mathscr{I}(\mu_\alpha)$, $\mathscr{I}(\mu_\beta)$ approach zero from one side, the disparity between the terms of any pair diminishes, and the change over in dominance is accomplished continuously for any y. Thus there is in fact no discontinuity in the functions at the crossing of the ξ axis, and the inversion theorem can be applied.

The same results could have been established by applying the inversion theorem with respect to η directly to the expressions (2.14), (2.15) and considering the contributions from the poles of the integrands in the η-plane for variations in $\mathscr{I}(\mu_\alpha)$, $\mathscr{I}(\mu_\beta)$.

4. Convergence of solution as $x \to \infty$

In obtaining conditions that the solution converge as $x \to \infty$, we bear in mind that the expressions for u and v may be different for different ranges of x, and that we shall be using those that obtain as $x \to \infty$. When the source is taken at $x = a$, we shall use expressions valid for $x > a$, noting that u and v for $x < a$ will need to be obtained by making certain changes in these.

We return to (2.14) and (2.15) and invert with respect to ξ, recovering x. The process parallels that of Section 3, except that $p^{\omega\xi}$ is a function of ξ and appears after inversion in a convolution integral. We introduce the notation

$$\eta^2+\kappa_\alpha^2 = \sigma_\alpha^2, \qquad \eta^2+\kappa_\beta^2 = \sigma_\beta^2, \tag{4.1}$$

and consider the case $\mathscr{I}(\sigma_\alpha) > 0$, $\mathscr{I}(\sigma_\beta) > 0$. Annihilating the coefficients of $\exp(-i\sigma_\alpha x)$ and $\exp(-i\sigma_\beta x)$ as $x \to \infty$, we obtain from the convergence of $u^{\omega\eta}(x, \eta, \omega)$ the two relations

$$\int_0^\infty \exp(i\sigma_\alpha\tau)\{2i\eta\sigma_\alpha u^\omega(\tau, 0, \omega) + (\kappa_\beta^2 - 2\kappa_\alpha^2 - 2\eta^2)v^\omega(\tau, 0, \omega)\}\, d\tau +$$

$$+ (\kappa_\beta^2 + 2\eta^2)u_\eta + 2i\eta\sigma_\alpha v_\eta = \frac{\eta}{\rho\beta^2}P(\sigma_\alpha, \omega), \tag{4.2}$$

$$\int_0^\infty \exp(i\sigma_\beta\tau)\{(\kappa_\beta^2+2\eta^2)u^\omega(\tau,0,\omega)+2i\eta\sigma_\beta v^\omega(\tau,0,\omega)\}\,d\tau -$$

$$-2i\eta\sigma_\beta u_\eta+(\kappa_\beta^2+2\eta^2)v_\eta = -\frac{i\sigma_\beta}{\rho\beta^2}P(\sigma_\beta,\omega), \tag{4.3}$$

where

$$P(\sigma,\omega) = \int_0^\infty p^\omega(\tau,\omega)\exp(i\sigma\tau)\,d\tau. \tag{4.4}$$

The same relations could have been obtained as necessary conditions for convergence of $v^{\omega\eta}(x,\eta,\omega)$ as $x\to\infty$.

Solving for u_η and v_η we get

$$F(\eta)u^{\omega\eta}(0,\eta,\omega) = \frac{\eta}{\rho\beta^2}\{(\kappa_\beta^2+2\eta^2)P(\sigma_\alpha,\omega)-2\sigma_\alpha\sigma_\beta P(\sigma_\beta,\omega)\}-$$

$$-(\kappa_\beta^2+2\eta^2)J_\alpha+2i\eta\sigma_\alpha J_\beta, \tag{4.5}$$

$$F(\eta)v^{\omega\eta}(0,\eta,\omega) = \frac{i\sigma_\beta}{\rho\beta^2}\{2\eta^2P(\sigma_\alpha,\omega)-(\kappa_\beta^2+2\eta^2)P(\sigma_\beta,\omega)\}- \tag{4.6}$$

$$-2i\eta\sigma_\beta J_\alpha-(\kappa_\beta^2+2\eta^2)J_\beta,$$

where

$$F(\eta) = (\kappa_\beta^2+2\eta^2)^2-4\eta^2\sigma_\alpha\sigma_\beta,$$

$$J_\alpha = \int_0^\infty \exp(i\sigma_\alpha\tau)\{2i\eta\sigma_\alpha u^\omega(\tau,0,\omega)+(\kappa_\beta^2-2\kappa_\alpha^2-2\eta^2)v^\omega(\tau,0,\omega)\}\,d\tau, \tag{4.7}$$

$$J_\beta = \int_0^\infty \exp(i\sigma_\beta\tau)\{(\kappa_\beta^2+2\eta^2)u^\omega(\tau,0,\omega)+2i\eta\sigma_\beta v^\omega(\tau,0,\omega)\}\,d\tau. \tag{4.8}$$

Here again, the sign of $i\sigma_\alpha$, $i\sigma_\beta$ is to be changed when $\mathscr{I}(\sigma_\alpha)<0$, $\mathscr{I}(\sigma_\beta)<0$.

5. The initial impulse

Now suppose that

$$p(x,t) = -Q\delta(x-a)\Phi(t), \tag{5.1}$$

where Q and a are constants and $\delta(\ )$ is the Dirac delta-function. This represents a normal pressure on the quarter-space, along the line $x=a$. The notation has been chosen to agree with that of Lamb (1904).

Then

$$p^\omega(x,\omega) = -\frac{1}{\sqrt{(2\pi)}}\int_{-\infty}^\infty Q\delta(x-a)\Phi(t)\exp(i\omega t)\,dt$$

$$= -Q\delta(x-a)\phi(\omega) \tag{5.2}$$

where

$$\phi(\omega) = \frac{1}{\sqrt{(2\pi)}}\int_{-\infty}^\infty \Phi(t)\exp(i\omega t)\,dt, \tag{5.3}$$

the Fourier transform of $\Phi(t)$.

Again

$$p^{\omega\xi} = -\int_0^\infty Q\delta(x-a)\phi(\omega)\exp(-\xi x)\,dx$$

$$= -Q\exp(-a\xi)\phi(\omega), \tag{5.4}$$

and

$$P(\sigma_\alpha,\omega) = -\int_0^\infty Q\delta(\tau-a)\phi(\omega)\exp(i\sigma_\alpha\tau)\,d\tau$$

$$= -Q\exp(i\sigma_\alpha a)\phi(\omega),$$

$$P(\sigma_\beta,\omega) = -Q\exp(i\sigma_\beta a)\phi(\omega). \tag{5.5}$$

The expressions in (5.5) are functions of η and ω.

6. The first approximation on $y = 0$

The transforms of the two components of displacement on $y = 0$ are $u^{\omega\xi}(\xi, 0, \omega)$ and $v^{\omega\xi}(\xi, 0, \omega)$. They are given by (3.12) and (3.13), which contain direct contributions from the source and also contributions from the integrals I_α, I_β containing the transforms $u^\omega(0, \nu, \omega)$, $v^\omega(0, \nu, \omega)$ of the displacements on $x = 0$.

We obtain our first approximation (in operational form $u_1^{\omega\xi}(\xi, 0, \omega)$, $v_1^{\omega\xi}(\xi, 0, \omega)$) for the disturbance on $y = 0$ by disregarding the terms containing I_α, I_β. We are, in effect, disregarding the boundary $x = 0$, and we obtain the solution for a half-space instead of quarter-space. Substituting $p^{\omega\xi}$ from (5.4) we have

$$u_1^{\omega\xi}(\xi,0,\omega) = \frac{\xi(\kappa_\beta^2+2\xi^2-2\mu_\alpha\mu_\beta)}{F(\xi)}\frac{Q}{\mu}\exp(-a\xi)\phi(\omega),$$

$$v_1^{\omega\xi}(\xi,0,\omega) = \frac{i\mu_\alpha \operatorname{sgn}\mathscr{I}(\mu_\alpha)\kappa_\beta^2}{F(\xi)}\frac{Q}{\mu}\exp(-a\xi)\phi(\omega). \tag{6.1}$$

Inverting to recover x and t we have

$$u_1(x,0,t) = \frac{1}{\sqrt{(2\pi)}}\int_{-\infty}^{\infty}\exp(-i\omega t)\phi(\omega)\,d\omega \times$$

$$\times\frac{1}{2\pi i}\int_{Br}\frac{\xi(\kappa_\beta^2+2\xi^2-2\mu_\alpha\mu_\beta)}{F(\xi)}\frac{Q}{\mu}\exp(-a\xi+x\xi)\,d\xi,$$

$$v_1(x,0,t) = \frac{1}{\sqrt{(2\pi)}}\int_{-\infty}^{\infty}\exp(-i\omega t)\phi(\omega)\,d\omega \times$$

$$\times\frac{1}{2\pi i}\int_{Br}\frac{i\mu_\alpha \operatorname{sgn}\mathscr{I}(\mu_\alpha)\kappa_\beta^2}{F(\xi)}\frac{Q}{\mu}\exp(-a\xi+x\xi)\,d\xi. \tag{6.2}$$

These expressions refer to the region $x > a$ and form the response to the impulse of time variation

$$\Phi(t) = \frac{1}{\sqrt{(2\pi)}} \int_{-\infty}^{\infty} \exp(-i\omega t)\phi(\omega)\, d\omega. \tag{6.3}$$

In order to compare this with Lamb's solution for the surface disturbance due to a concentrated normal pressure of harmonic time-variation on a half-space, we write the response to a line-force of time variation $\exp(-i\omega t)$ and put $a = 0$, obtaining

$$u_1 = -\frac{iQ}{2\pi\mu} \int_{Br} \frac{\xi(\kappa_\beta^2 + 2\xi^2 - 2\mu_\alpha\mu_\beta)}{F(\xi)} \exp(\xi x)\, d\xi \exp(-i\omega t),$$

$$v_1 = \frac{Q}{2\pi\mu} \int_{Br} \frac{\mu_\alpha \operatorname{sgn} \mathscr{I}(\mu_\alpha)\kappa_\beta^2}{F(\xi)} \exp(\xi x)\, d\xi \exp(-i\omega t). \tag{6.4}$$

Following Lamb, we take ω real and fixed. Then the singularities of the integrand are branch points at $\xi = \pm i\kappa_\alpha, \pm i\kappa_\beta$ and poles at $\pm i\kappa_\gamma$, where $\kappa_\gamma = \omega/\gamma$, γ being the Rayleigh-wave velocity corresponding to the relevant pair of roots of $F(\xi) = 0$. Since there are no singularities of the integrand in the positive half of the ξ-plane we can take the Bromwich contour to be the imaginary axis, indented to avoid singularities. Then writing $\xi = i\zeta$ we get from (6.4)

$$\left.\begin{aligned} u_1 &= -\frac{iQ}{2\pi\mu} \int_{-\infty}^{\infty} \frac{\zeta(2\zeta^2 - \kappa_\beta^2 - 2\lambda_\alpha\lambda_\beta)}{F(\zeta)} \exp(i\zeta x - i\omega t)\, d\zeta, \\ v_1 &= -\frac{Q}{2\pi\mu} \int_{-\infty}^{\infty} \frac{\lambda_\alpha \operatorname{sgn}\zeta\, \kappa_\beta^2}{F(\zeta)} \exp(i\zeta x - i\omega t)\, d\zeta, \end{aligned}\right\} \text{ for } x > 0, \tag{6.5}$$

where

$$\lambda_\alpha = \sqrt{(\zeta^2 - \kappa_\alpha^2)}, \qquad \lambda_\beta = \sqrt{(\zeta^2 - \kappa_\beta^2)},$$

and

$$F(\zeta) = (2\zeta^2 - \kappa_\beta^2)^2 - 4\zeta^2\lambda_\alpha\lambda_\beta.$$

(6.5) are the same as the expressions given by Lamb (1904, equations (52)), if account is taken of Lamb's addition of a free surface wave to his original solution.

In order to ascertain the shape and size of the Rayleigh-wave which moves out from the source, we revert to (6.4), restore the factor $\exp(-a\xi)$, and distort the contour in the ξ-plane. These expressions are valid for the region $x > a$, and thus the integrands vanish exponentially as $|\xi| \to \infty$ in the region $\mathscr{R}(\xi) < 0$. If ω is positive, we swing the upper part of the contour so that the path of integration takes the form shown in Figure 2(a). If ω is negative, we swing the lower part of the contour to the position shown in Figure 2(b). Then the Rayleigh wave

corresponding to source of time variation $\exp(-i\omega t)$ is given by $2\pi i$ times the residue at $\xi = i\kappa_\gamma = i\omega/\gamma$:

$$\left.\begin{aligned} u_{xR} &= \frac{Q}{\mu}\frac{i\kappa_\gamma}{F'(i\kappa_\gamma)}(\kappa_\beta^2 - 2\kappa_\gamma^2 + 2\kappa_{\alpha\gamma}\kappa_{\beta\gamma})\exp\{i\kappa_\gamma(x-a) - i\omega t\}, \\ v_{xR} &= \pm\frac{iQ}{\mu}\frac{i\kappa_{\alpha\gamma}}{F'(i\kappa_\gamma)}\kappa_\beta^2\exp\{i\kappa_\gamma(x-a) - i\omega t\}, \end{aligned}\right\} \tag{6.6}$$

where

$$\kappa_{\alpha\gamma} = \sqrt{(\kappa_\gamma^2 - \kappa_\alpha^2)}, \qquad \kappa_{\beta\gamma} = \sqrt{(\kappa_\gamma^2 - \kappa_\beta^2)}, \tag{6.7}$$

and the upper or lower sign is taken according as $\omega \gtrless 0$ (since the value $\mu_\alpha = i\kappa_{\alpha\gamma}$ gives the correct sign for $\mathcal{I}(\mu_\alpha)$ whether $\omega \gtrless 0$).

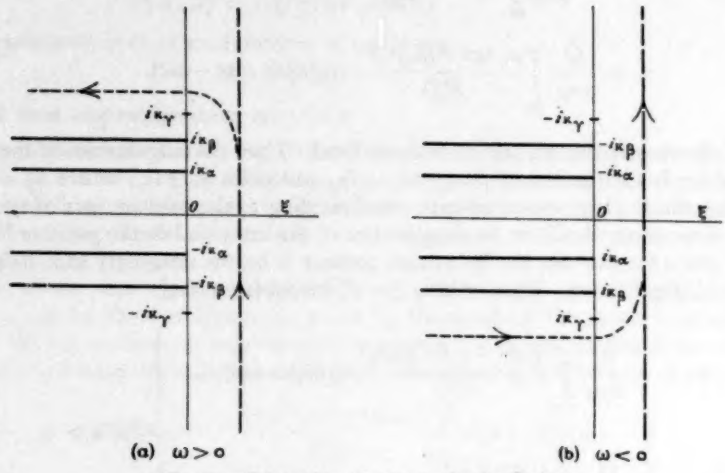


FIG. 2.—Distortion of contour in the ξ-plane.
(a) for $\omega > 0$ (b) for $\omega < 0$.

Now from (3.14)

$$F'(\xi) = \xi\left\{8(\kappa_\beta^2 + 2\xi^2) - 8\mu_\alpha\mu_\beta - 4\xi^2\left(\frac{\mu_\beta}{\mu_\alpha} + \frac{\mu_\alpha}{\mu_\beta}\right)\right\},$$

and so

$$\begin{aligned} \frac{F'(i\kappa_\gamma)}{i\kappa_\gamma} &= 8(\kappa_\beta^2 - 2\kappa_\gamma^2) + 8\kappa_{\alpha\gamma}\kappa_{\beta\gamma} + 4\kappa_\gamma^2\left(\frac{\kappa_{\beta\gamma}}{\kappa_{\alpha\gamma}} + \frac{\kappa_{\alpha\gamma}}{\kappa_{\beta\gamma}}\right) \\ &= \omega^2 G \text{ (say)}, \end{aligned} \tag{6.8}$$

where G is a quantity depending only on α and β. Hence the wave moving off to $x = +\infty$ has

$$\left.\begin{aligned} u_{xR} &= \frac{Q}{\mu\omega^2 G}(\kappa_\beta^2 - 2\kappa_\gamma^2 + 2\kappa_{\alpha\gamma}\kappa_{\beta\gamma})\exp[-i\omega\{t - (x-a)/\gamma\}] \\ v_{xR} &= -\frac{Q}{\mu\omega^2 G}\frac{\kappa_{\alpha\gamma}\kappa_\beta^2}{\kappa_\gamma}\left(\frac{1}{i}\operatorname{sgn}\omega\right)\exp[-i\omega\{t - (x-a)/\gamma\}]. \end{aligned}\right\} \tag{6.9}$$

These expressions are the same as those found by Lamb (1904, equations (73) and (76)).

7. The allied functions $\Phi(t)$ and $\Phi'(t)$

The factor $\pm(1/i)$, corresponding to $\omega \gtrless 0$, in the expression for v_{xR} in (6.9), is of great significance. For while the use of (6.3) enables us quickly to derive from the first equation of (6.2) the result

$$u_{1R}(x,0,t) = \frac{Q}{\mu G}(\beta^{-2} - 2\gamma^{-2} + 2\gamma_\alpha^{-1}\gamma_\beta^{-1})\Phi\left(t - \frac{x-a}{\gamma}\right), \tag{7.1}$$

where

$$\gamma_\alpha^{-2} = \gamma^{-2} - \alpha^{-2}, \qquad \gamma_\beta^{-2} = \gamma^{-2} - \beta^{-2},$$

we obtain from the second equation

$$v_{1R}(x,0,t) = -\frac{Q}{\mu G}\frac{\gamma}{\beta^2\gamma_\alpha}\frac{1}{\sqrt{(2\pi)}}\int_{-\infty}^{\infty}\frac{\operatorname{sgn}\omega}{i}\phi(\omega)\exp(-i\omega t)\,d\omega. \tag{7.2}$$

If in the integral in (7.2) we insert $\phi(\omega)$ from (5.3) we obtain

$$\begin{aligned}&\frac{1}{2\pi}\int_{-\infty}^{\infty}\exp(-i\omega t)\left(\frac{\operatorname{sgn}\omega}{i}\right)\int_{-\infty}^{\infty}\Phi(\sigma)\exp(i\omega\sigma)\,d\sigma d\omega\\ &\qquad = \frac{1}{\pi}\int_{-\infty}^{\infty}d\omega\int_{-\infty}^{\infty}\Phi(\sigma)\sin\omega(\sigma-t)\,d\sigma.\end{aligned} \tag{7.3}$$

This function of t is the Hilbert transform of $\Phi(t)$, or the *allied function* of $\Phi(t)$ (Jeffreys & Jeffreys 1956, Section 14.111), and denoted (following Lamb, who used it first) by $\Phi'(t)$. Thus

$$\Phi(t) = \frac{1}{\sqrt{(2\pi)}}\int_{-\infty}^{\infty}\phi(\omega)\exp(-i\omega t)\,d\omega = \frac{1}{\pi}\int_{0}^{\infty}d\omega\int_{-\infty}^{\infty}\Phi(\sigma)\cos\omega(\sigma-t)\,d\sigma, \tag{7.4}$$

$$\Phi'(t) = \frac{1}{\sqrt{(2\pi)}}\int_{-\infty}^{\infty}\frac{\operatorname{sgn}\omega}{i}\phi(\omega)\exp(-i\omega t)\,d\omega = \frac{1}{\pi}\int_{0}^{\infty}d\omega\int_{-\infty}^{\infty}\Phi(\sigma)\sin\omega(\sigma-t)\,d\sigma. \tag{7.5}$$

Examples of allied functions are given by Lamb (1904, equations (102), (103), (104)), Titchmarsh (1937, Ch. V), Arons & Yennie (1950) and Jeffreys & Lapwood (1957).

Thus the Rayleigh pulse travelling along $x > a$ is given as to $u_{1R}(x,0,t)$ by (7.1) and as to $v_{1R}(x,0,t)$ by

$$v_{1R}(x,0,t) = -\frac{Q}{\mu G}\frac{\gamma}{\beta^2\gamma_\alpha}\Phi'\left(t - \frac{x-a}{\gamma}\right). \tag{7.6}$$

The difference in shape of u_{1R} and v_{1R} for given x may be very considerable. For instance, if

$$\Phi(t) = \frac{\tau}{t^2+\tau^2},$$

then

$$\Phi'(t) = -\frac{t}{t^2+\tau^2},$$

and if

$$\Phi(t) = H(t) \quad \text{(Heaviside unit function)},$$

then

$$\Phi'(t) = -\frac{1}{\pi}\ln|t-t_1| \quad \text{(Jeffreys 1940)}.$$

In the latter case $\Phi(t)$ has a finite step, whereas $\Phi'(t)$ has a logarithmic infinity, and non-zero values before $t = 0$. This ideal case will not be found in actual phenomena, but it is to be expected that $\Phi'(t)$ will show a gradual beginning and large amplitude corresponding to $\Phi(t)$ with a sudden beginning and sharp initial rise.

8. The first approximations on $x = 0$

We next consider the approximation for the disturbance on $x = 0$ which is obtained by neglecting in (4.5) and (4.6) the terms containing J_α and J_β. We have then only the immediate effect of the source. Substituting from (5.5) for $P(\sigma_\alpha, \omega)$ and $P(\sigma_\beta, \omega)$ we find, for $\mathscr{I}(\sigma_\alpha, \sigma_\beta) > 0$,

$$\left.\begin{aligned} u_1^{\omega\eta}(0,\eta,\omega) &= -\frac{Q}{\mu}\phi(\omega)\frac{\eta}{F(\eta)}\{(\kappa_\beta^2+2\eta^2)\exp(i\sigma_\alpha a)-2\sigma_\alpha\sigma_\beta\exp(i\sigma_\beta a)\},\\ v_1^{\omega\eta}(0,\eta,\omega) &= -\frac{Q}{\mu}\phi(\omega)\frac{i\sigma_\beta}{F(\eta)}\{2\eta^2\exp(i\sigma_\alpha a)-(\kappa_\beta^2+2\eta^2)\exp(i\sigma_\beta a)\}. \end{aligned}\right\} \quad (8.1)$$

If $\mathscr{I}(\sigma_\alpha, \sigma_\beta) < 0$, the quantities $i\sigma_\alpha$, $i\sigma_\beta$ are replaced in (8.1) by $-i\sigma_\alpha$, $-i\sigma_\beta$. Then

$$\begin{pmatrix} u_1 \\ v_1 \end{pmatrix}(0,y,t) = \frac{1}{\sqrt{(2\pi)}}\int_{-\infty}^{\infty}\exp(-i\omega t)\,d\omega\,\frac{1}{2\pi i}\int_{Br}\begin{pmatrix} u_1 \\ v_1 \end{pmatrix}^{\omega\eta}(0,\eta,\omega)\exp(\eta y)\,d\eta. \quad (8.2)$$

The response to $\exp(-i\omega t)$ is

$$\begin{aligned} u_1(0,y,t) &= -\frac{1}{2\pi i}\times\\ &\times\frac{Q}{\mu}\int_{Br}\frac{\eta}{F(\eta)}\{(\kappa_\beta^2+2\eta^2)\exp(\eta y \pm i\sigma_\alpha a - i\omega t)-2\sigma_\alpha\sigma_\beta\exp(\eta y \pm i\sigma_\beta a - i\omega t)\}\,d\eta,\\ v_1(0,y,t) &= -\frac{1}{2\pi i}\times\\ &\times\frac{Q}{\mu}\int_{Br}\frac{(\pm i\sigma_\beta)}{F(\eta)}\{2\eta^2\exp(\eta y \pm i\sigma_\alpha a - i\omega t)-(\kappa_\beta^2+2\eta^2)\exp(\eta y \pm i\sigma_\beta a - i\omega t)\}\,d\eta, \end{aligned} \quad (8.3)$$

where the upper or lower sign is chosen according as $\mathscr{I}(\sigma_\alpha, \sigma_\beta) \gtrless 0$. If we evaluate the integral by distortion of the path in the η-plane, we obtain the displacement due to the direct P-wave from the contribution to the integral from the path in the neighbourhood of the branch point $\eta = i\kappa_\alpha$, and that due to the direct S-wave from the neighbourhood of $\eta = i\kappa_\beta$. We find the earliest Rayleigh-wave at $(0, y)$ by evaluating the contribution to (8.3) from the loop around the pole $\eta = i\kappa_\gamma$. Hence the expression for the residue will differ according as ω is positive or negative, and we obtain the factor

$$\begin{aligned}&\exp\{i\kappa_\gamma y \pm ia\sqrt{(\kappa_\alpha^2 - \kappa_\gamma^2)} - i\omega t\} \\ &\quad = \exp\left\{\mp\frac{a\omega}{\gamma_\alpha} - i\omega\left(t - \frac{y}{\gamma}\right)\right\}, \text{ according as } \omega \gtrless 0,\end{aligned} \tag{8.4}$$

in the first term, and the factor

$$\exp\left\{\mp\frac{a\omega}{\gamma_\beta} - i\omega\left(t - \frac{y}{\gamma}\right)\right\}, \text{ according as } \omega \gtrless 0, \tag{8.5}$$

in the second term. These may be interpreted as follows.

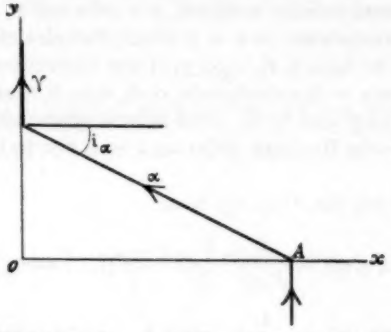


FIG. 3.—Rayleigh waves generated on Oy by body-waves from A.

In Figure 3 is shown the path of a ray which travels with velocity α from the source to impinge on $x = 0$ at an angle with the normal of $i_\alpha = \sin^{-1}(\gamma/\alpha)$ and then travels with velocity γ along $y = 0$. The time of travel is

$$\begin{aligned}&(a\sec i_\alpha)/\alpha + (y - a\tan i_\alpha)/\gamma \\ &\quad = y/\gamma + a\sec i_\alpha\{1/\alpha - (\sin i_\alpha)/\gamma\} \\ &\quad = y/\gamma.\end{aligned} \tag{8.6}$$

Similarly the time of travel of a ray which travels with velocity β across the medium and impinges at angle $\sin^{-1}(\gamma/\beta)$ is y/γ. Thus we can interpret (8.5), which seems to be a Rayleigh wave starting from the corner at $t = 0$, as the Rayleigh-wave set up on $x = 0$ by incident P- and S-waves, as found by Lamb (1904) and Nakano (1925). The factor $\exp(\mp a\omega/\gamma_{\alpha,\beta})$ as $\omega \gtrless 0$ shows that this wave decreases exponentially with increasing a. It will therefore be very small (for large a) compared with the wave established on $x = 0$ by the Rayleigh wave along $y = 0$, which will travel with amplitude independent of a. It is this latter wave which we now investigate.

9. The second approximation on $x = 0$, including the transmitted Rayleigh pulse

The expressions given in (6.2) for $u_1(x, 0, t)$, $v_1(x, 0, t)$ refer to the region $x > a$. Since there is symmetry about $x = a$ for the initial pulse, we can construct by inspection the complete specification on $y = 0$ and write, in operational form as far as t is concerned,

$$\left.\begin{aligned} u_1{}^{\omega}(x, 0, \omega) &= \mp \frac{Q}{\mu}\phi(\omega)\frac{1}{2\pi i}\int_{Br} \frac{\xi(\kappa_\beta{}^2 + 2\xi^2 - 2\mu_\alpha\mu_\beta)}{F(\xi)}\exp\{\pm(a-x)\xi\}\,d\xi, \\ v_1{}^{\omega}(x, 0, \omega) &= \frac{Q}{\mu}\phi(\omega)\frac{1}{2\pi i}\int_{Br} \frac{i\mu_\alpha\kappa_\beta{}^2 \operatorname{sgn}\mathscr{I}(\mu_\alpha)}{F(\xi)}\exp\{\pm(a-x)\xi\}\,d\xi, \end{aligned}\right\} \tag{9.1}$$

where the upper or lower sign is taken according as $x \lessgtr a$.

In a qualitative evaluation of the results so far gained, we may say that we have in (9.1) a representation of the disturbance on $y = 0$ when pulses have set out from the source at $(a, 0)$ and before the earliest pulse strikes the boundary at $x = 0$. In (8.1) we have, in operational form, the disturbance which arises when P- and S-pulses, as if in a semi-infinite medium, are reflected at the boundary $x = 0$. To find a second approximation at $x = 0$ which includes effects of the boundary $y = 0$, we must add to $u_1(0, y, t)$, $v_1(0, y, t)$ the corrections obtained when the first approximations on $y = 0$, namely $u_1(x, 0, t)$, $v_1(x, 0, t)$ are substituted into the integrals J_α and J_β of (4.5) and (4.6). This second approximation will include the pulse we are seeking—the Rayleigh pulse on $x = 0$ due to the Rayleigh pulse on $y = 0$.

These corrections are, for $\mathscr{I}(\sigma_\alpha, \sigma_\beta) > 0$,

$$\left.\begin{aligned} u_2{}^{\omega\eta}(0, \eta, \omega) &= \frac{1}{F(\eta)}\{-(\kappa_\beta{}^2 + 2\eta^2)J_{\alpha 1} + 2i\eta\sigma_\alpha J_{\beta 1}\}, \\ v_2{}^{\omega\eta}(0, \eta, \omega) &= \frac{1}{F(\eta)}\{-2i\eta\sigma_\beta J_{\alpha 1} - (\kappa_\beta{}^2 + 2\eta^2)J_{\beta 1}\}, \end{aligned}\right\} \tag{9.2}$$

where

$$\left.\begin{aligned} J_{\alpha 1} &= 2i\eta\sigma_\alpha\mu_{\alpha 1} + (\kappa_\beta{}^2 - 2\kappa_\alpha{}^2 - 2\eta^2)v_{\alpha 1}, \\ J_{\beta 1} &= (\kappa_\beta{}^2 + 2\eta^2)u_{\beta 1} + 2i\eta\sigma_\beta v_{\beta 1}, \end{aligned}\right\} \tag{9.3}$$

with

$$\left.\begin{aligned} u_{\alpha 1} &= \int_0^\infty u_1{}^{\omega}(\tau, 0, \omega)\exp(i\sigma_\alpha\tau)\,d\tau, \\ u_{\beta 1} &= \int_0^\infty u_1{}^{\omega}(\tau, 0, \omega)\exp(i\sigma_\beta\tau)\,d\tau, \end{aligned}\right\} \tag{9.4}$$

and $v_{\alpha 1}$, $v_{\beta 1}$ similarly defined.

Since we want to find the part of the second approximation on $x = 0$ which is due to Rayleigh pulses on $y = 0$, we evaluate the contributions to the integrals

in (9.1) from the loops around the Rayleigh pole $\xi = i\kappa_\gamma$, and as in (6.6) we get, for $\omega > 0$

$$\left.\begin{aligned} u_R^{\omega}(\tau, 0, \omega) &= \mp\frac{Q}{\mu}\phi(\omega)\frac{i\kappa_\gamma}{F'(i\kappa_\gamma)}(\kappa_\beta^2 - 2\kappa_\gamma^2 + 2\kappa_{\alpha\gamma}\kappa_{\beta\gamma})\exp\{\pm i\kappa_\gamma(a-\tau)\}, \\ v_R^{\omega}(\tau, 0, \omega) &= \frac{iQ}{\mu}\phi(\omega)\frac{i\kappa_{\alpha\gamma}}{F'(i\kappa_\gamma)}\kappa_\beta^2\exp\{\pm i\kappa_\gamma(a-\tau)\}, \end{aligned}\right\} \tag{9.5}$$

according as $\tau \lesseqgtr a$. When $\omega < 0$, the expression for u_R^{ω} is unchanged, but in v_R^{ω} the factor $i\kappa_{\alpha\gamma}$ is replaced by $(-i\kappa_{\alpha\gamma})$. We introduce new constants

$$\left.\begin{aligned} H &= \frac{Q}{\mu}\frac{i\kappa_\gamma}{F'(i\kappa_\gamma)}(\kappa_\beta^2 - 2\kappa_\gamma^2 + 2\kappa_{\alpha\gamma}\kappa_{\beta\gamma}) = \frac{Q}{\mu}\frac{\beta^{-2} - 2\gamma^{-2} + 2\gamma_\alpha^{-1}\gamma_\beta^{-1}}{G}, \\ K &= \frac{Q}{\mu}\frac{i\kappa_{\alpha\gamma}}{F'(i\kappa_\gamma)}\kappa_\beta^2 = \frac{Q}{\mu}\frac{\gamma}{\beta^2\gamma_\alpha G}. \end{aligned}\right\} \tag{9.6}$$

Then substituting from (9.5), (9.6) into (9.4) we obtain those parts of $u_{\alpha 1}$ and $u_{\beta 1}$ which are due to the primary Rayleigh pulse on $y = 0$:

$$\left.\begin{aligned} u_{\alpha R} &= H\phi(\omega)\times \\ &\quad\times\Big[-\int_0^a \exp\{i(\sigma_\alpha - \kappa_\gamma)\tau + i\kappa_\gamma a\}\,d\tau + \int_a^\infty \exp\{i(\sigma_\alpha + \kappa_\gamma)\tau - i\kappa_\gamma a\}\,d\tau\Big], \\ v_{\alpha R} &= \pm iK\phi(\omega)\times \\ &\quad\times\Big[\int_0^a \exp\{i(\sigma_\alpha - \kappa_\gamma)\tau + i\kappa_\gamma a\}\,d\tau + \int_a^\infty \exp\{i(\sigma_\alpha + \kappa_\gamma)\tau - i\kappa_\gamma a\}\,d\tau,\Big] \end{aligned}\right\} \tag{9.7}$$

where the sign is $\pm$ according as $\omega \gtrless 0$.

If these expressions for $u_{\alpha R}$, $v_{\alpha R}$ are substituted into (9.3) and the consequent value of $J_{\alpha 1}$ put into (9.1), $u_2^{\omega\eta}(0, \eta, \omega)$ will contain a term of the type $\psi(\eta)/F(\eta)u_{\alpha R}$. When inverted with respect to η and ω this will give, as part of $u_2(0, y, t)$, a triple integral in which the exponential factor of the integrand is

$$\begin{aligned} &\exp\{-i\omega t + \eta y + i\kappa_\gamma(a-\tau) + i\sigma_\alpha\tau\} \\ &\qquad = \exp\{-i\omega t + \eta y + i\omega(a-\tau)/\gamma + i\omega(1 + \alpha^2\eta^2/\omega^2)^{\frac{1}{2}}\tau/\alpha\}. \end{aligned}$$

If in this we write $\alpha\eta = i\omega\sin\theta$, and hence $(1 + \alpha^2\eta^2/\omega^2)^{\frac{1}{2}} = \cos\theta$, we get

$$\exp\left\{-i\omega\left(t - \frac{y\sin\theta}{\alpha} - \frac{\tau\sin\theta}{\alpha} - \frac{a-\tau}{\gamma}\right)\right\}. \tag{9.8}$$

This characterizes an element of the disturbance which has travelled (see Figure 4) along AT with speed γ (i.e. as a Rayleigh pulse) and along TB with speed α. The first integral in (9.7) sums such arrivals over values of τ between 0 and a; the second similarly sums arrivals from points on Ox beyond A, these also having been excited by the arriving Rayleigh pulse.

The convergence of the integrals in (9.7) is ensured by the fact that $\mathscr{I}(\sigma_\alpha) > 0$. Evaluating them, we get, for $\mathscr{I}(\sigma_\alpha) > 0$,

$$\left.\begin{aligned} u_{\alpha R} &= H\phi(\omega)\left(-\frac{i\exp(i\kappa_\gamma a)}{\sigma_\alpha-\kappa_\gamma}+\frac{2i\sigma_\alpha\exp(i\sigma_\alpha a)}{\sigma_\alpha{}^2-\kappa_\gamma{}^2}\right), \\ v_{\alpha R} &= \pm K\phi(\omega)\left(-\frac{\exp(i\kappa_\gamma a)}{\sigma_\alpha-\kappa_\gamma}+\frac{2\kappa_\gamma\exp(i\sigma_\alpha a)}{\sigma_\alpha{}^2-\kappa_\gamma{}^2}\right) \text{ as } \omega \gtrless 0. \end{aligned}\right\} \tag{9.9}$$

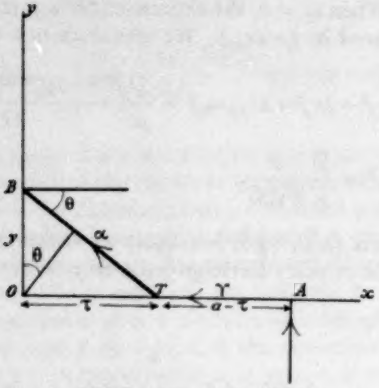


FIG. 4.—The path *ATB*.

$u_{\beta R}$ and $v_{\beta R}$ are given by similar expressions.

Similarly, when $\mathscr{I}(\sigma_\alpha) < 0$, $i\sigma_\alpha$ in (9.7) is everywhere replaced by $-i\sigma_\alpha$ and the integrals converge. Then corresponding to (9.9) we have expressions in which $-i\sigma_\alpha$ replaces $i\sigma_\alpha$.

First consider the second term of the expression for $u_{\alpha R}$ in (9.9). This will contribute through $J_{\alpha 1}$ to $u_2{}^{\omega\eta}(0, \eta, \omega)$ a term containing the factor $\exp(\pm i\sigma_\alpha a)$ according as $\mathscr{I}(\sigma_\alpha) \gtrless 0$. When we invert with respect to η to obtain $u_2{}^{\omega}(0, y, \omega)$ we have an integral along the Bromwich contour in the η-plane. The contribution to this integral (after distortion of the contour) from the Rayleigh pole $\eta = i\kappa_\gamma$ will contain the factor $\exp(\mp \omega a/\gamma_\alpha)$ according as $\omega \gtrless 0$. This diminishes exponentially as a increases, and so the expression under consideration is not part of the transmitted Rayleigh pulse we are seeking. Similarly, contributions from the neighbourhoods of the branch points are exponentially small for large a. The only other possible singularities of this second term are at $\sigma_\alpha = \pm\kappa_\gamma$. If $\omega > 0$ the singularity of the second term at $\eta = \omega/\gamma_\alpha$ is cancelled by that of the first term, and the singularity at $\eta = -\omega/\gamma_\alpha$ is not near to the path of integration in original or distorted form. Similarly, if $\omega < 0$ the singularity at $\eta = -\omega/\gamma_\alpha$ is cancelled and that at $\eta = \omega/\gamma_\alpha$ is far from the path of integration.

We therefore consider only the first terms in $u_{\alpha R}$, etc., to get

$$u_{\alpha R'} = \frac{-iH\phi(\omega)\exp(i\kappa_\gamma a)}{\sigma_\alpha-\kappa_\gamma}, \qquad v_{\alpha R'} = \frac{-K\phi(\omega)\exp(i\kappa_\gamma a)}{\sigma_\alpha-\kappa_\gamma}, \tag{9.10}$$

with like terms for $u_{\beta R'}$, $v_{\beta R'}$ β replacing α in (9.10). We may note that by evaluating at $\eta = i\kappa_\gamma$ we have picked out from the phases in (9.8) that for which $\sin\theta = \alpha/\gamma > 1$, apparently an imaginary angle of incidence. It is better to say that we

have picked not a plane wave but a surface wave, in fact the one which travels with Rayleigh velocity.

From (9.2), (9.3) and (9.10) we obtain, for $\omega \gtrless 0$,

$$\left.\begin{aligned} u_{2R}{}^{\omega\eta}(0,\eta,\omega) &= \frac{\exp(i\kappa_\gamma a)\phi(\omega)}{F(\eta)}\times \\ &\times\left[-\frac{\kappa_\beta{}^2+2\eta^2}{\pm\sigma_\alpha-\kappa_\gamma}\{\pm 2\eta\sigma_\alpha H\pm(2\eta^2+2\kappa_\alpha{}^2-\kappa_\beta{}^2)K\}\pm\right. \\ &\left.\pm\frac{2i\eta\sigma_\alpha}{\pm\sigma_\beta-\kappa_\gamma}\{-i(\kappa_\beta{}^2+2\eta^2)H-2i\eta\sigma_\beta K\}\right], \\ v_{2R}{}^{\omega\eta}(0,\eta,\omega) &= \frac{\exp(i\kappa_\gamma a)\phi(\omega)}{F(\eta)}\times \\ &\times\left[\mp\frac{2i\eta\sigma_\beta}{\mp\sigma_\alpha-\kappa_\gamma}\{\pm 2\eta\sigma_\alpha H\pm(2\eta^2+2\kappa_\alpha{}^2-\kappa_\beta{}^2)K\}-\right. \\ &\left.-\frac{\kappa_\beta{}^2+2\eta^2}{\pm\sigma_\beta-\kappa_\gamma}\{-i(\kappa_\beta{}^2+2\eta^2)H-2i\eta\sigma_\beta K\}\right]. \end{aligned}\right\} \tag{9.11}$$

To find the Rayleigh wave along $x = 0$ which is part of (9.11), we invert with respect to η and evaluate the contribution from the pole at $\eta = i\kappa_\gamma$, obtaining, after simplification,

$$\left.\begin{aligned} u^{\omega}{}_{2RR'} &= \frac{Q}{\mu}\frac{2\gamma^{-2}-\beta^{-2}}{2\gamma^{-2}-\alpha^{-2}}L\left(\frac{\gamma}{\gamma_\alpha}\pm\frac{1}{i}\right)\phi(\omega)\exp[-i\omega\{t-(a+y)/\gamma\}], \\ v^{\omega}{}_{2RR'} &= \frac{Q}{\mu}\frac{2\gamma^{-1}\gamma_\beta{}^{-1}}{2\gamma^{-2}-\alpha^{-2}}L\left(1\mp\frac{1}{i}\frac{\gamma}{\gamma_\alpha}\right)\phi(\omega)\exp[-i\omega\{t-(a+y)/\gamma\}], \end{aligned}\right\} \tag{9.12}$$

where

$$L = 2G^{-2}\gamma\gamma_\alpha{}^{-1}\beta^{-2}(\beta^{-2}-\alpha^{-2}). \tag{9.13}$$

Finally we invert with respect to ω and express the result in terms of Φ and Φ' as:

$$\left.\begin{aligned} u_{2RR'}(0,y,t) &= \frac{Q}{\mu}\frac{2\gamma^{-2}-\beta^{-2}}{2\gamma^{-2}-\alpha^{-2}}L\left\{\frac{\gamma}{\gamma_\alpha}\Phi(\tau)+\Phi'(\tau)\right\}, \\ v_{2RR'}(0,y,t) &= \frac{Q}{\mu}\frac{2\gamma^{-1}\gamma_\beta{}^{-1}}{2\gamma^{-2}-\alpha^{-2}}L\left\{\Phi(\tau)-\frac{\gamma}{\gamma_\alpha}\Phi'(\tau)\right\}, \end{aligned}\right\} \tag{9.14}$$

where

$$\tau = t-(a+y)/\gamma. \tag{9.15}$$

10. Further approximations

The method of iteration here begun could be continued, so that at any stage the approximations already found for $u(0, y, t)$, $v(0, y, t)$ can be substituted into (3.15), (3.16) to give, by way of (3.12), (3.13), the next approximation on $y = 0$. Similarly an approximation to the displacements on $y = 0$ will give the next approximation, by way of (4.7), (4.8) and (4.5), (4.6), on $x = 0$.

In following out this procedure, we would in fact be tracing the growth of the disturbance. At any stage, points of the boundary $y = 0$ will be centres of disturbances radiating in all directions into the medium. The parts which induce

further disturbance on $x = 0$ are summed in the integrals (4.7), (4.8). From the exceedingly complex sequence we can pick out, by physical argument, the main features. There will be P and S pulses radiated from the source at A, and these will give rise to reflected pulses PP, PS, SP, SS from the boundary $x = 0$. None of these, viewed from the standpoint of ray theory, will produce further reflections as they move off into the medium. There will be also those long-wave effects which are not included in ray theory. But in so far as these are, for example, Rayleigh waves on $x = 0$ due to the impinging of body waves from the source directly on $x = 0$, or Rayleigh waves on $y = 0$ due to radiation from points on $x = 0$, they will contain exponential factors like those discussed in Section 8 or those arising from the second term in (9.8). Some of these may attain importance in the immediate neighbourhood of the corner, but at greater distances only two of these diffraction terms will be important—that discussed in Section 9 which represents a pulse travelling all along the boundary as a surface wave and transmitted round the corner, and the Rayleigh pulse reflected at the corner back along $y = 0$. In this two-dimensional case each of these travels unchanged in shape or amplitude once it has undergone transformation at the corner. It is hoped to discuss the reflected pulse in a future paper. Meanwhile the changes in pulse shape described by (9.14) are more fully explored with the help of an example.

11. Comparison of incident and transmitted Rayleigh pulses

Collecting results, we see that the Rayleigh pulse which travels along the face $y = 0$ towards the corner has displacements

$$\left.\begin{aligned} u_{1R}(x,0,t) &= \frac{Q}{\mu G}(2\gamma^{-2}-\beta^{-2}-2\gamma_\alpha^{-1}\gamma_\beta^{-1})\Phi\left(t-\frac{a-x}{\gamma}\right), \\ v_{1R}(x,0,t) &= -\frac{Q}{\mu G}\frac{\gamma}{\beta^2\gamma_\alpha}\Phi'\left(t-\frac{a-x}{\gamma}\right). \end{aligned}\right\} \tag{11.1}$$

(These are derived from (7.1) and (7.6) by making the changes consequent on taking $x-a < 0$.)

The transmitted Rayleigh pulse, which travels along the face $x = 0$ from the corner, has displacements (9.14)

$$\left.\begin{aligned} u_{2RR'}(0,y,t) &= \frac{QL}{\mu}\frac{2\gamma^{-2}-\beta^{-2}}{2\gamma^{-2}-\alpha^{-2}}\left\{\frac{\gamma}{\gamma_\alpha}\Phi\left(t-\frac{a+y}{\gamma}\right)+\Phi'\left(t-\frac{a+y}{\gamma}\right)\right\}, \\ v_{2RR'}(0,y,t) &= \frac{QL}{\mu}\frac{2\gamma^{-1}\gamma_\alpha^{-1}}{2\gamma^{-2}-\alpha^{-2}}\left\{\Phi\left(t-\frac{a+y}{\gamma}\right)-\frac{\gamma}{\gamma_\alpha}\Phi'\left(t-\frac{a+y}{\gamma}\right)\right\}. \end{aligned}\right\} \tag{11.2}$$

As an illustration, we compute the numerical coefficients for the Poisson material in which $\lambda = \mu$, $\alpha = \sqrt{3}\beta$. Then $\gamma = 0{\cdot}9194\beta$, and we find, for the incident pulse

$$\left.\begin{aligned} u_{1R} &= 0{\cdot}1250\frac{Q}{\mu}\Phi\left(t-\frac{a-x}{\gamma}\right), \\ v_{1R} &= -0{\cdot}1835\frac{Q}{\mu}\Phi'\left(t-\frac{a-x}{\gamma}\right), \end{aligned}\right\} \tag{11.3}$$

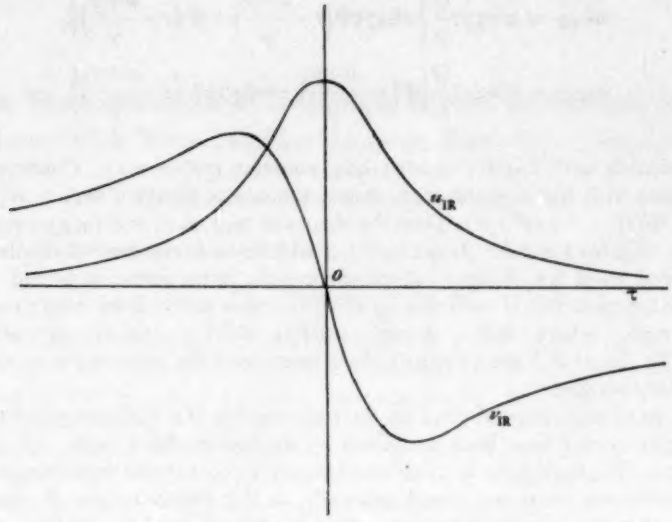


FIG. 5.—Incident Rayleigh pulse.

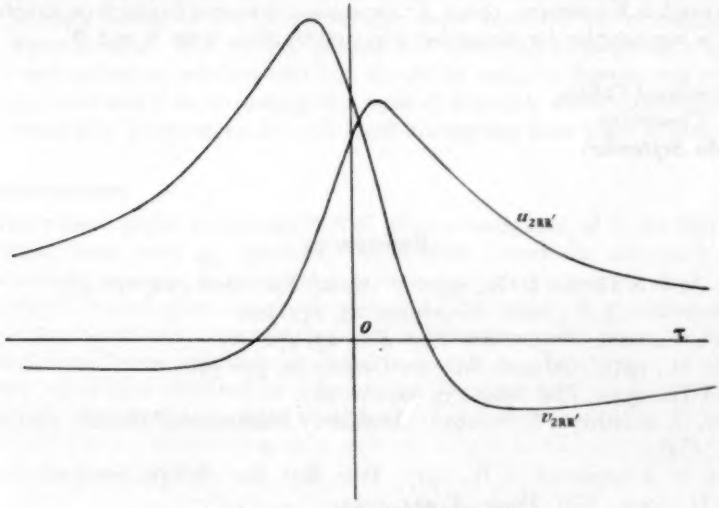


FIG. 6.—Transmitted Rayleigh pulse.
(Vertical scale 4 × that of Figure 5.)

and for the transmitted pulse

$$u_{2RR'} = 0{\cdot}3357\frac{Q}{\mu}\left\{0{\cdot}8475\Phi\left(t-\frac{a+y}{\gamma}\right)+\Phi'\left(t-\frac{a+y}{\gamma}\right)\right\},$$
$$v_{2RR'} = 0{\cdot}0242\frac{Q}{\mu}\left\{\Phi\left(t-\frac{a+y}{\gamma}\right)-0{\cdot}8475\Phi'\left(t-\frac{a+y}{\gamma}\right)\right\}. \qquad (11.4)$$

(11.3) coincide with Lamb's result (1904, equation (78) *et seq.*). Continuing the comparison with his work, we now choose the source function $\Phi(t) = \tau/(\tau^2+t^2)$, so that* $\Phi'(t) = -t/(\tau^2+t^2)$. Then the shapes of (u_{1R}, v_{1R}) and $(u_{2RR'}, v_{2RR'})$ are shown in Figures 5 and 6.† It is clear that whichever component of displacement is measured, there is a change of shape of the pulse as the corner is turned.

If the function $\Phi(t)$ is such that its allied function differs from it very markedly (for example, where $\Phi(t) = A\exp(-\sigma t)H(t)$, $\Phi'(t) = -(A/\pi)\exp(-\sigma t)Ei(\sigma t)$, as taken by Arons & Yennie (1950)), the distortion of the pulse at the corner may also be very marked.

Records of experimental work on the transmission of a Rayleigh pulse round a right-angled corner have been published by de Bremaecker (1958). They show clearly that a Rayleigh pulse is transmitted round the corner and does change shape. The experiments were not aimed primarily at the determination of change of shape, so that the contribution of the allied function is hard to identify. de Bremaecker considered that a phase change of $\frac{1}{4}\pi$ takes place at the corner, but his graphs may bear other interpretations. I am indebted to Dr Menzel for sending me records of a similar experiment performed at the Institut für Geophysik (Rao, Gutdeutsch & Klussmann, 1959). In those also a distorted Rayleigh pulse is found, but it is too complex for separation into contributions from Φ and Φ'.

Emmanuel College,
Cambridge.
1960 *September.*

References

Arons, A. B. & Yennie, D. R., 1950. *J. Acoust. Soc. Amer.*, **22**, 231–237.
de Bremaecker, J. C., 1958. *Geophysics*, **23**, 253–266.
Jeffreys, H., 1926. *Proc. Camb. Phil. Soc.*, **23**, 472–481.
Jeffreys, H., 1931. *Gerlands Beit. zur Geoph.*, **30**, 336–350.
Jeffreys, H., 1940. *Phil. Mag.* (7), **30**, 161–167.
Jeffreys, H. & Jeffreys, B. S., 1956. *Methods of Mathematical Physics.* (3rd ed) (C.U.P.)
Jeffreys, H. & Lapwood, E. R., 1957. *Proc. Roy. Soc. A*, **241**, 455–479.
Lamb, H., 1904. *Phil. Trans. A*, **203**, 1–42.
Nakano, H., 1925. *Jap. J. Astr. Geophys.*, **2**, 233–326.
Rao, S. H., Gutdeutsch, R. & Klussmann, J., 1959. *Zeit. für Geophys.*, **25**, 253–264.
Titchmarsh, E. C., 1937. *Theory of Fourier Integrals.* (O.U.P.)

* The modern definition of the allied function differs in sign from Lamb's.

† I am grateful to P. R. Lapwood for help in computations and the drawing of figures.

The Transmission of Rayleigh Waves across an Ocean Floor with Two Surface Layers, Part II: Numerical

Urs Hochstrasser and Robert Stoneley

"Under the water it rumbled on."—*S. T. Coleridge, The Ancient Mariner*

Summary

This investigation obtains values of the wave velocity and group velocity of waves of Rayleigh type, following ocean paths corresponding to the structures indicated by experimental investigations using explosions at sea. Eight different models have been considered, with ocean depths 4 or 6 km, a sedimentary layer of depth 1 or 2 km, underlain by basic rocks of thickness 5 or 7 km; in all cases these structures were supposed to rest on ultrabasic rocks of great thickness.

The wave velocity was found for a series of values of the wave-number by solving an eleven-row determinantal equation; the computations were performed by the use of the SEAC electronic computer of the U.S. National Bureau of Standards, and from these values the corresponding values of the group velocity were obtained by numerical differentiation. The general agreement of these dispersion curves with observed values indicates that they should be useful in determining the ray track and time of passage of a train of Rayleigh waves crossing an ocean floor whose structure and depth are varying from place to place.

1. Introduction

The present paper constitutes Part II of an investigation of which Part I was published three years ago (Stoneley 1957). Part I gave the derivation of the wave velocity equation for periodic disturbances moving across an ocean of uniform depth H resting on a layer of sediments, of thickness T_1, beneath which is a layer, in general basic, of rock of thickness T_2, extending down to the Mohorovičić discontinuity, below which the material is denoted by the suffix 3. The wave velocity equation is obtained as a determinantal equation of the eleventh order, which, for an assigned structure, can be solved numerically to give values of the wave velocity c corresponding to a series of values of the parameter κ, where $2\pi/\kappa$ is the wave length.

2. Calculation of the wave velocity

To agree with the layering of the ocean floor, so far as it is known, the values $H = 4, 6$ km; $T_1 = 1, 2$ km; $T_2 = 5, 7$ km have been chosen, corresponding to eight different possibilities, so that in any actual basin of extent much greater than the wave length of a surface wave linear interpolation (and probably extrapolation) would be permissible. The velocity of sound in sea-water is c_0, and the

velocities of compressional and distortional waves in the elastic solid media are written as α, β, with appropriate suffixes. The density is denoted by ρ. The constants used in the computations are

Ocean:	$H = 4, 6$ km;	$c_0 = 1{\cdot}53$ km/s,		$\rho_0 = 1{\cdot}03$ g/cm^3
Sediments	$T_1 = 1, 2$ km;	$\alpha_1 = 1{\cdot}8$ km/s,	$\beta_1 = 1{\cdot}0$ km/s,	$\rho_1 = 1{\cdot}8$ g/cm^3
Basic	$T_2 = 5, 7$ km;	$\alpha_2 = 6{\cdot}5$ km/s,	$\beta_2 = 3{\cdot}7$ km/s,	$\rho_2 = 2{\cdot}9$ g/cm^3
Ultrabasic		$\alpha_3 = 8{\cdot}1$ km/s,	$\beta_3 = 4{\cdot}5$ km/s,	$\rho_3 = 3{\cdot}4$ g/cm^3

The solution of the wave velocity equation was effected by use of the SEAC electronic computer of the U.S. National Bureau of Standards, Washington, D.C. A sequence of values of the parameter κT_2 was assumed, and the equation solved for c/β_2. For values of κT_2 greater than about 4 the accuracy of the solution became doubtful. The solution is not complete, in the sense that only the fundamental type of oscillation was obtained; some values of higher modes were in fact calculated, but a systematic evaluation of the many possible types of oscillation that may exist was not contemplated in the present research.

From the values of the wave velocity c the values of group velocity C were calculated by numerical differentiation according to the formula $C = c + \kappa\, dc/d\kappa$. This process is a very unsatisfactory one; it involves considerable loss of accuracy, and where both a forward difference formula and a central difference formula can be used on the same set of values, spaced at equal values of the argument, the discrepancy is notable. Thus the values of C quoted in the tables of this paper are not trustworthy to the third decimal place, particularly for the larger values of κT_2.

The period τ can be found from the formula

$$\tau = \frac{2\pi}{\kappa c} = \frac{2\pi T_2}{\beta_2} \cdot \frac{1}{\kappa T_2 \,.\, c/\beta_2}$$

and is not affected by the inaccuracies of numerical differentiation. The values of c/β_2, τ and C calculated for the eight assumed structures are listed in the following tables.

$H = 6$, $T_1 = 1$, $T_2 = 7$

κT_2	c/β_2	τ(s)	C(km/s)
0·0	1·12347	∞	4·157
0·1	1·11090	76·4	4·064
0·2	1·09886	38·6	3·975
0·3	1·08637	26·0	3·874
0·4	1·07163	19·8	3·713
0·5	1·05075	16·2	3·393
0·6	1·01527	13·9	2·708
0·7	0·95651	12·7	1·707
0·8	0·88529	12·0	1·206
0·9	0·81838	11·5	1·017
1·0	0·76110	11·2	0·834
1·2	0·67321	10·5	0·874
1·4	0·61128	9·9	0·910
1·6	0·56641	9·4	0·959
1·8	0·53298	8·9	1·014
2·0	0·50744	8·4	1·045
3·0	0·43934	6·4	1·087
4·0	0·41044	5·2	?

$H = 6$, $T_1 = 2$, $T_2 = 7$

κT_2	c/β_2	τ(s)	C(km/s)
0·0	1·12347	∞	4·157
0·1	1·10906	76·6	4·051
0·2	1·09486	38·8	3·946
0·3	1·07959	26·2	3·813
0·4	1·06075	20·0	3·603
0·5	1·03253	16·4	3·153
0·6	0·98340	14·4	2·188
0·7	0·91073	13·3	1·372
0·8	0·83460	12·7	0·938
0·9	0·76792	12·3	0·862
1·0	0·71232	11·9	0·803
1·2	0·62801	11·3	0·778
1·4	0·56875	10·7	0·809
1·6	0·52562	10·0	0·835
1·8	0·49320	9·6	0·892
2·0	0·46810	9·1	0·905
3·0	0·39672	7·1	?
4·0	0·35900	5·9	(1·019)

$H = 6$
$T_1 = 1$
$T_2 = 5$

κT_2	c/β_2	τ(s)	C(km/s)
0·0	1·12347	∞	4·157
0·1	1·10827	76·6	4·048
0·2	1·09239	38·9	3·903
0·3	1·07219	26·4	3·705
0·4	1·03662	20·5	3·027
0·5	0·96144	17·7	1·861
0·6	0·86183	16·4	1·044
0·7	0·77545	15·6	0·870
0·8	0·70778	15·0	0·856
0·9	0·65513	14·4	0·880
1·0	0·61368	13·8	0·903
1·2	0·55377	12·8	0·977
1·4	0·51351	11·8	1·038
1·6	0·48518	10·9	1·087
1·8	0·46449	10·2	1·115
2·0	0·44887	9·5	1·148
3·0	0·40661	7·0	?
4·0	0·38512	5·5	?

$H = 6$
$T_1 = 2$
$T_2 = 5$

κT_2	c/β_2	τ(s)	C(km/s)
0·0	1·12347	∞	4·157
0·1	1·10564	76·8	4·022
0·2	1·08630	39·1	3·865
0·3	1·06028	26·7	3·578
0·4	1·01098	21·0	2·682
0·5	0·91529	18·6	1·255
0·6	0·81045	17·5	0·854
0·7	0·72589	16·7	0·777(?)
0·8	0·66086	16·1	0·780
0·9	0·61050	15·2	0·738
1·0	0·57084	14·9	0·802
1·2	0/51319	13·8	0·864
1·4	0·47387	12·8	0·902
1·6	0·44556	11·9	0·910
1·8	0·42417	11·1	0·954
2·0	0·40729	10·4	?
3·0	0·35295	8·0	?
4·0	0·31570	6·7	?

$H = 4$
$T_1 = 1$
$T_2 = 7$

κT_2	c/β_2	τ(s)	C(km/s)
0·0	1·12347	∞	4·157
0·1	1·11240	76·3	4·077
0·2	1·10227	38·5	4·006
0·3	1·09265	25·9	3·937
0·4	1·08303	19·6	3·864
0·5	1·07267	15·8	3·764
0·6	1·06048	13·3	3·622
0·7	1·04472	11·6	3·390
0·8	1·02260	10·4	2·997
0·9	0·99067	9·5	2·398
1·0	0·94810	9·0	1·723
1·2	0·85128	8·3	1·042
1·4	0·76620	7·9	0·864
1·6	0·69890	7·6	0·826
1·8	0·64620	7·3	0·844
2·0	0·60452	7·0	0·851
3·0	0·48693	5·8	0·861
4·0	0·43499	4·9	(0·877)

$H = 4$
$T_1 = 2$
$T_2 = 7$

κT_2	c/β_2	τ(s)	C(km/s)
0·0	1·12347	∞	4·157
0·1	1·11058	76·4	4·063
0·2	1·09835	38·7	3·975
0·3	1·08628	26·1	3·885
0·4	1·07362	19·8	3·775
0·5	1·05925	16·0	3·629
0·6	1·04137	13·6	3·399
0·7	1·01695	11·9	3·009
0·8	0·98177	10·8	2·382
0·9	0·93404	10·1	1·682
1·0	0·87919	9·7	1·261
1·2	0·77523	9·1	0·773
1·4	0·69247	8·8	0·681
1·6	0·62874	8·4	0·671
1·8	0·57913	8·1	0·649
2·0	0·53981	7·9	0·723
3·0	0·42577	6·6	?
4·0	0·37001	5·7	?

	$H = 4$ $T_1 = 1$ $T_2 = 5$				$H = 4$ $T_1 = 2$ $T_2 = 5$		
κT_2	c/β_2	τ(s)	C(km/s)	κT_2	c/β_2	τ(s)	C(km/s)
0·0	1·12347	∞	4·157	0·0	1·12347	∞	4·157
0·1	1·11048	76·5	4·063	0·1	1·10787	76·6	4·041
0·2	1·09801	38·7	3·974	0·2	1·09224	38·9	3·915
0·3	1·08519	26·1	3·861	0·3	1·07518	26·3	3·756
0·4	1·07046	19·8	3·649	0·4	1·05407	20·1	3·535
0·5	1·05063	16·2	3·434	0·5	1·02316	16·6	3·044
0·6	1·01888	13·9	2·840	0·6	0·97178	14·6	2·175
0·7	0·96630	12·5	1·966	0·7	0·89815	13·5	1·307
0·8	0·89826	11·8	1·264	0·8	0·82131	12·9	0·873
0·9	0·83111	11·4	1·104	0·9	0·75374	12·5	0·799
1·0	0·77230	11·0	0·847	1·0	0·69713	12·2	0·694
1·2	0·68064	10·4	0·819	1·2	0·61067	11·6	0·658
1·4	0·61527	9·9	0·843	1·4	0·54925	11·0	0·651
1·6	0·56750	9·4	0·884	1·6	0·50402	10·5	0·710
1·8	0·53160	8·9	0·928	1·8	0·48402	9 7	?
2·0	0·50395	8·4	0·972	2·0	0·44244	9·6	?
3·0	0·42821	6·6	?	3·0	0·36191	7·8	?
4·0	0·39337	5·4	?	4·0	0·31754	6·7	?

3. Comparison with observations

A number of attempts have been made by various authors (e.g. Stoneley 1926; Wilson & Baykal 1948; Ewing & Press 1952) to compare the observed values of the dispersion curves (i.e. C *vs* τ) with those calculated according to an assumed constitution of the ocean and its underlying layers. Evidently the assumption of uniform layers is an oversimplification, and will yield only an overall comparison. Berckhemer in 1956 considered the effect of variation in water depth on the time of transmission; it was, in fact, along these lines that the present investigation was

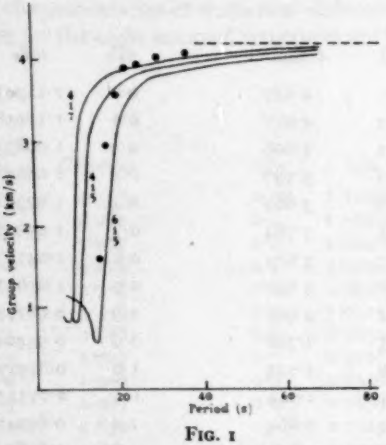


FIG. 1

planned, so that not only might an overall comparison be made, but one in which variation of oceanic structure could be allowed for (remembering, of course, that the variation in surface wave velocity determines the departure of the path from a great circle route).

Three of the dispersion curves are shown in Figure 1. The seven black dots on the diagram represent readings taken from a paper by Ewing & Press (1952)

relating to the earthquake of 1950 July 29 in the Solomon Islands. While the general fit is good, the readings do not fit closely any one of the three curves, and the same is true of a comparison with the remaining calculated curves, to the extent that these observations do not directly decide which is the better *average* constitution of the ocean and floor. Some adjustments of the adopted values of the density and elastic constants may be indicated by future researches. Figure 2 shows the observations of Wilson & Baykal and of Berckhemer compared with four of the calculated dispersion curves. The two sets of observations, of course, are not comparable, although both relate to the Atlantic Ocean, but the fit is sufficiently good to justify further attempts to reconcile the observed dispersion curves of Rayleigh waves with the present knowledge of the ocean floor.

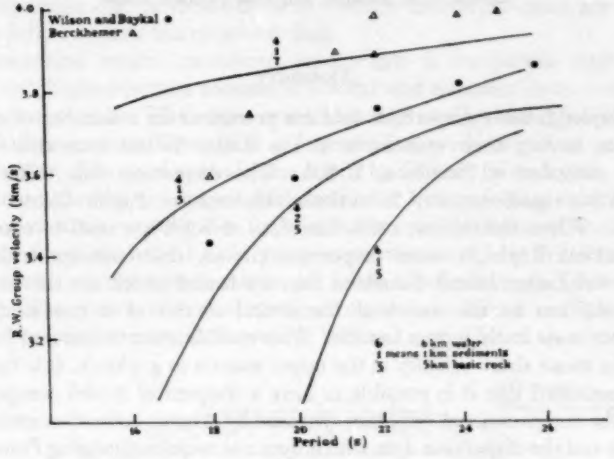


FIG. 2

In investigations, such as those of J. and M. Darbyshire (1956), on the transmission of the waves of microseisms from mid-ocean to a coastal station, the tables given here will facilitate the computation of the ray path and the time of transit.

This work, in part, was carried out at the U.S. National Bureau of Standards, Washington, D.C., under a grant from the U.S. Office of Naval Research. We wish to express our indebtedness to the Director of the National Bu au of Standards for the facilities afforded to us, and for permission to publish this paper.

16 *Millington Road*,
Cambridge.
1960 *October*.

National Bureau of Standards,
Washington, D.C.,
U.S.A.

References

Darbyshire, J. & M., 1956. *Mon. Not. R. Astr. Soc. Geophys. Suppl.*, **7**, 301.
Ewing, M. & Press, F., 1952. *Bull. Seismol. Soc. Amer.*, **42**, 315.
Stoneley, R., 1926. *Mon. Not. R. Astr. Soc. Geophys. Suppl.*, **1**, 349.
Stoneley, R., 1957. *Bull. Seismol. Soc. Amer.*, **47**, 7.
Wilson, J. T. & Baykal, O., 1948. *Bull. Seismol. Soc. Amer.*, **38**, 46.

Rayleigh Wave Dispersion and Crustal Structure in the Eastern Pacific and Indian Oceans*

Robert L. Kovach and Frank Press

"A wave has beat from that distant shore to this, and now sinks in the unknown depths."—*E. Bramah, Kai Lung's Golden Hours*

Summary

Rayleigh wave dispersion data are presented for a number of earthquakes having their epicentres in the Easter Island area and which were recorded at Pasadena, U.S.A. The dispersion data differ in a small but significant way from those obtained for Pacific Ocean basin paths. When the seismic refraction data of Raitt are used to compute theoretical Rayleigh wave dispersion curves, discrepancies with the observed Easter Island–Pasadena data are found which are resolved by the addition to the assumed theoretical model of a modified low velocity zone in the upper mantle. This modification consists of lowering the mean shear velocity in the upper mantle to 4·5 km/s. It is further demonstrated that it is possible to have a theoretical model compatible with the compressional velocities obtained by seismic refraction measurements and the dispersion data which does not require changing Poisson's ratio from the usual value and consequent reduction of the shear velocities in the upper mantle. This is accomplished by reducing the thickness of the layer with compressional velocity 8·2 km/s to the minimal thickness required to give a refraction arrival. The mean shear velocity of 4·5 km/s implies a "soft" upper mantle along the Easter Island to Pasadena path. Dispersion data also show that the mean thickness of unconsolidated sediments between Easter Island and Peru is 0·57 km.

Rayleigh wave dispersion data are also presented for Indian Ocean shocks recorded at Wilkes Station, Antarctica. Within the limits of observational error, the Indian Ocean data are in agreement in the period range from 37 to 25 seconds with the theoretical model assumed for the Easter Island to Pasadena data.

1. Introduction

The utilization of Rayleigh wave dispersion data for the study of oceanic areas has been used previously in many regions (Oliver, Ewing & Press 1955). However, when Rayleigh wave dispersion data are used alone to determine crustal structure a unique solution cannot usually be given since many different velocity

* Contribution No. 993, Division of Geological Sciences, California Institute of Technology.

depth functions can give a satisfactory fit to the dispersion data. In addition, in the study of oceanic areas care must be used in any determination of crustal structure because of the effect of the water layer on the dispersion of Rayleigh waves. Press (1960) and Oliver, Kovach & Dorman (1960) have shown earlier in the study of continental areas that it is possible to determine a more unique crustal structure when surface wave dispersion data are coupled with seismic refraction data.

In this paper Rayleigh wave dispersion data in the period range from about 35 to 12 seconds are presented for an oceanic path from the Easter Island area to Pasadena, U.S.A. With the aid of the constraints presented by the seismic refraction data of Raitt (1956), theoretical Rayleigh wave dispersion curves are computed in an attempt to fit the observed data. Since the technique of combining refraction and dispersion data is relatively new, various theoretical cases are calculated in order to fully examine the observed data.

The theoretical model considered here which is compatible both with the refraction and dispersion data consists of a water and sediment layer overlying two rock layers in the crust. The upper crustal layer ($\beta = 2{\cdot}93$ km/s) is 1 km in thickness and the lower crustal layer ($\beta = 3{\cdot}87$ km/s) is 4 km in thickness overlying the mantle. Further data on this model are listed in Table 3 (Case 1588).

The second part of this paper deals with the Indian Ocean, which is one of the great areas of the world today for which few studies of crustal structure have been made. As part of the I.G.Y. programme the California Institute of Technology operated a seismograph station in Antarctica from 1957 April 17 to 1959 January 31. The station is located at 66° 20′ south latitude and 110° 31′ east longitude in the region of Antarctica known as Wilkes Land (Figure 1). As can be seen

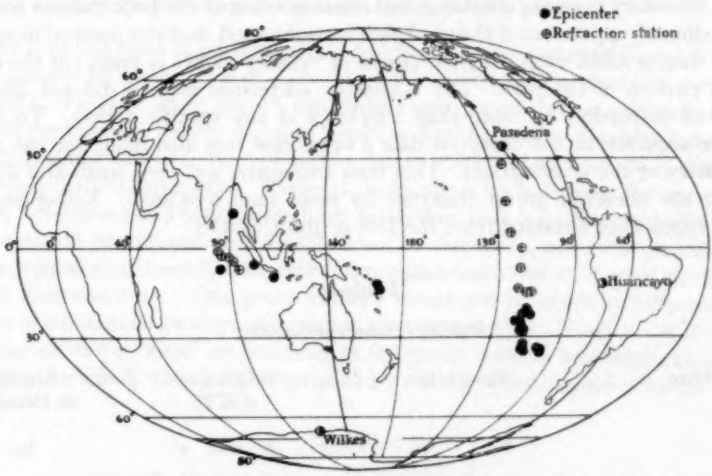


FIG. 1.—Index map showing earthquake epicentres and refraction data in the Pacific and Indian Oceans.

from Figure 1 the seismograph station at Wilkes is ideally located for the recording of earthquakes which have their epicentres in the Indian Ocean. The number of earthquakes selected for the study of the Indian Ocean is limited, however, because of the scarcity of shocks having a long oceanic path to Wilkes. Within

the limits of the experimental data the Indian Ocean data show the same maximum group velocity as that observed for the Easter Island to Pasadena data.

2. Data and discussion

Figure 1 is an index map of the epicentres, seismograph stations, and seismic refraction stations utilized in this study. The following symbols are used:

α = compressional wave velocity
β = shear wave velocity
ρ = density
H = layer thickness
k = wave number
C = phase velocity
U = group velocity
T = wave period
T_0 = free period of seismometer
T_g = free period of galvanometer

Table 1 is a list of the origin times and epicentral distances for the Easter Island shocks recorded at Pasadena. The shocks were recorded on the long-period Benioff seismograph ($T_0 = 1$ s, $T_g = 90$ s). As a means of obtaining precision, earthquakes were selected which had approximately the same great circle path to Pasadena. This multiplicity of data gives increased confidence in reading dispersion by the conventional graphical method. A correction to the observed dispersion was applied for the small non-oceanic portion of the great circle path. If the boundary between continent and ocean is taken at the 1 000 fathom contour, approximately 4 per cent of the total path is continental, and any correction applied to the data is small even if a poor choice of velocity values is made for the continental portion of the path. The maximum correction applied did not alter the observed dispersion by more than 0·05 km/s at any specific period. To obtain further accuracy in the observed data a correction was also made for the group time delay of the seismograph. This time correction was very small and did not change the observed group velocities by more than 0·01 km/s. Epicentres and origin times were obtained from the ISS or the USCGS.

Table 1

Pacific Ocean earthquakes

Date	Epicentre		Origin time (GCT)			Distance to Pasadena
	S	W	h	min	s	km
1940 January 2	28½°	113½°	11	07	18	6953
1943 September 19	28½	113½	04	47	46	6953
1944 April 19	27	113	22	32	08	6791
1944 April 22	27	113	03	35	45	6791
1944 November 18	28½	113½	07	53	14	6953
1951 February 19	25	117	22	11	54	6548
1952 July 18	23	114½	18	39	40	6338
1960 April 15	27	113	03	25	38	6791

The resulting observed Rayleigh wave dispersion data are shown in Figure 2. As can be seen from Figure 2 the observed data are remarkably consistent for the number of shocks studied and practically no scatter is present. However, before an attempt is made to fit a theoretical dispersion curve to the observed data it is well at this point to discuss a few of the constraints or restrictions which must be placed on the theoretical model selected.

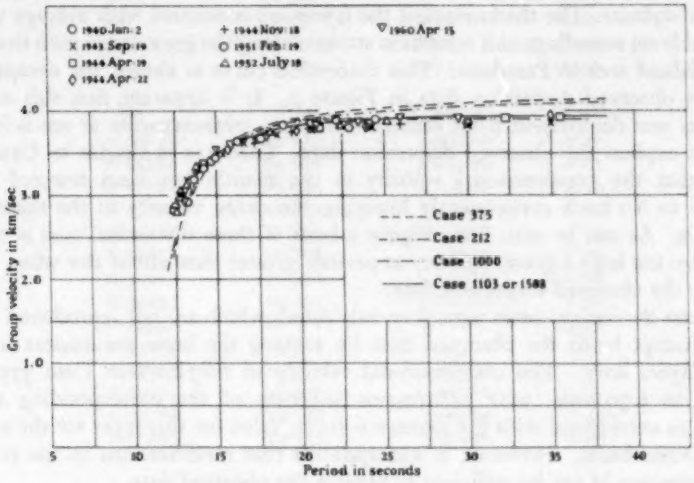


FIG. 2.—Experimental Rayleigh wave data, Easter Island to Pasadena, with theoretical dispersion curves for four models.

The first and most important restriction placed on the theoretical model is that of the seismic refraction data of Raitt (1956) for the Pacific Ocean which are summarized in Table 2. Even though refraction measurements at sea primarily determine compressional wave velocities, whereas Rayleigh wave dispersion data are more dependent on shear wave velocities, one is still faced with selecting a reasonable value for Poisson's ratio. However, since shear arrivals from the mantle are usually never measured at sea a possibility exists that there is a difference in Poisson's ratio between continental and oceanic mantle. Secondly, one must require that the theoretical model satisfies the mean maximum group velocity of 3·92 km/s shown by the observed data. This group velocity maximum is primarily influenced by the distribution of shear wave velocities in the upper mantle. Finally, the observed data also exhibit a "knee" or flattening of the group velocity beginning at a period of about 18 s which must also be satisfied by the theoretical model selected.

Table 2

Pacific Ocean seismic refraction results (after Raitt)

Mean layer thickness (km)	α (km/s)
0·29	2·15
1·02	5·09
4·18	6·76
	8·25

As a first attempt to fit the observed dispersion data with a theoretical model several dispersion curves were calculated based on Raitt's refraction data. All the theoretical Rayleigh wave dispersion curves discussed below were computed on the Bendix G-15D electronic digital computer using the Haskell formulation. The layer parameters and computed results for all the models are listed in Table 3. The first model calulated is Case 375 which consists of a water layer and two crustal layers overlying the mantle. The mantle is taken to be homogeneous and semi-infinite. The thicknesses of the layers are consistent with average values obtained from soundings and refraction stations along the great circle path from the Easter Island area to Pasadena. This theoretical curve is shown for comparison with the observed dispersion data in Figure 2. It is apparent that this crustal model as was determined from seismic refraction measurements at sea is inadequate to explain the observed dispersion data. Case 212 is similar to Case 375 except that the compressional velocity in the mantle has been lowered from 8·2 km/s to 8·1 km/s consequently lowering the shear velocity in the mantle to 4·68 km/s. As can be seen from Figure 2 both of these theoretical curves calculated have too high a group velocity at periods greater than about 16 s when compared to the observed dispersion data.

Various theoretical cases were then calculated, which are not reproduced here, in an attempt to fit the observed data by varying the layer parameters in the crustal layers only. The compressional velocity of 6·83 km/s in Case 375 was lowered to 6·70 km/s, with subsequent lowering of the corresponding shear velocity, to correspond with the average velocity value for this layer for the entire Pacific Ocean basin. However, it was apparent that modifications in the crustal layers alone would not be sufficient to explain the observed data.

In order to fit the maximum group velocity observed, one is forced either drastically to reduce the shear velocity in the mantle in a simple model with a semi-infinite mantle, or to use a more refined model which includes a low-velocity zone in the upper mantle. As has been shown earlier by Dorman, Ewing & Oliver (1960) a theoretical model with a low velocity zone in the upper mantle is a better approximation to the oceanic mantle structure. Case 1 000 is a theoretical model based on the crustal refraction arrivals except that a low-velocity zone in the mantle has been added. In addition, the liquid layer has been divided into a water and an unconsolidated sediment layer. The mantle structure selected is based on Case 8 099 of Dorman, Ewing & Oliver (1960) for the Pacific Ocean (see Table 3) except that a shallower depth to the velocity reversal has been used than in Case 8 099. A shallower depth to the velocity reversal in the mantle is necessary because the upper mantle structure given in Case 8 099 gives a group velocity maximum of about 4·01 km/s which is too high when compared with the observed Easter Island data. Case 1 000 is the simplest modification of Case 8 099 involving a shallower low-velocity layer. This model is also shown for comparison with the observed data in Figure 2. Case 1 000 gives a good fit to the dispersion data at periods greater than about 25 s but the theoretical curve still deviates from the observed data by about 0·1 km/s at a period of about 20 s. This "knee" or flattening of the observed group velocity data at 20 s requires that the shear velocity in the upper mantle be lowered to 4·5 km/s. In Case 1 103 the compressional velocity in the upper mantle is the required 8·2 km/s and the corresponding shear velocity has been changed to 4·5 km/s. Case 1 103, as can be seen in Figure 2, is an excellent theoretical fit to the observed data. However, since the compressional arrival in the mantle (*Pn*) is fixed at 8·2 km/s from the refraction measurements a shear

Table 3

Layer parameters and computed results for theoretical models

Case 375

H (km)	α (km/s)	β (km/s)	ρ (g/cm^3)
4·3	1·52		1·03
1·0	5·07	2·93	2·67
4·16	6·83	3·94	3·00
	8·20	4·73	3·30

kH	T (s)	C (km/s)	U (km/s)
0·16637	38·155	4·25600	4·16
0·19384	32·866	4·24080	4·13
0·22082	28·954	4·22560	4·10
0·24706	25·972	4·21040	4·06
0·27232	23·649	4·19520	4·03
0·29637	21·808	4·18000	3·99
0·31907	20·330	4·16480	3·94
0·34033	19·131	4·14960	3·90
0·39532	16·652	4·10400	3·73
0·45090	14·819	4·04320	3·47
0·49224	13·782	3·98240	3·16
0·52475	13·128	3·92160	2·84
0·55186	12·680	3·86080	2·54
0·57563	12·351	3·80000	2·26

Case 212

H (km)	α (km/s)	β (km/s)	ρ (g/cm^3)
4·3	1·52		1·03
1·0	5·07	2·93	2·67
4·16	6·83	3·94	3·00
	8·10	4·68	3·30

kH	T (s)	C (km/s)	U (km/s)
0·11277	56·49	4·2408	4·18
0·17031	37·68	4·2104	4·12
0·22672	28·51	4·1800	4·05
0·27972	23·276	4·1496	3·98
0·32744	20·031	4·11920	3·90
0·36894	17·910	4·08880	3·80
0·45985	14·697	3·99760	3·41
0·51823	13·346	3·90640	2·93
0·56095	12·624	3·81520	2·48

Case 1 000

H (km)	α (km/s)	β (km/s)	ρ (g/cm^3)
4·2	1·52		1·03
0·2	2·10	0·50	2·10
1·0	5·07	2·93	2·67
4·0	6·70	3·87	3·00
36·0	7·82	4·61	3·34
173·0	8·17	4·30	3·4425
	8·49	4·60	3·5265

kH	*T* (s)	*C* (km/s)	*U* (km/s)
0·10864	60·991	3·98240	3·916
0·11420	58·067	3·97936	3·92
0·11952	55·513	3·97708	3·93
0·32883	20·182	3·97628	3·91
0·33608	19·755	3·97480	3·90
0·36277	18·3361	3·96720	3·845
0·39766	16·792	3·95200	3·74
0·40884	16·359	3·94561	3·70
0·44907	15·009	3·91521	3·525
0·50927	13·443	3·85472	2·99
0·54181	12·838	3·79392	2·67
0·56933	12·416	3·73312	2·39
0·59374	12·103	3·67232	2·13
0·61622	11·858	3·61152	1·74

Case 1 103

H (km)	α (km/s)	β (km/s)	ρ (g/cm^3)
4·1	1·52		1·03
0·2	2·10	0·5	2·10
1·0	5·07	2·93	2·67
4·0	6·70	3·87	3·00
36·0	8·20	4·50	3·34
173·0	8·17	4·30	3·4425
	8·49	4·60	3·5265

kH	*T* (s)	*C* (km/s)	*U* (km/s)
0·15037	43·098	3·97480	3·95
0·17797	36·443	3·97176	3·96
0·23385	27·767	3·96720	3·94
0·24040	27·016	3·96644	3·94
0·26130	24·875	3·96340	3·93
0·28316	22·976	3·95960	3·90
0·31362	20·784	3·95200	3·86
0·38567	17·032	3·92160	3·71
0·40993	16·087	3·90640	3·63
0·43022	15·388	3·89120	3·55
0·44797	14·836	3·87600	3·48
0·52194	13·040	3·78480	2·96

Case 8 099 (Dorman)

H (km)	α (km/s)	β (km/s)	ρ (g/cm^3)
5·0	1·52		1·03
1·0	2·10	1·00	2·10
5·0	6·41	3·70	2·84
49·0	7·82	4·61	3·34
160·0	8·17	4·30	3·4425
100·0	8·49	4·60	3·5265
90·0	8·81	4·80	3·6040
90·0	9·32	5·19	3·7650
100·0	9·97	5·49	4·0100
100·0	10·48	5·79	4·2300
100·0	10·85	6·03	4·4100
100·0	11·12	6·20	4·5450
100·0	11·33	6·31	4·6400
	11·49	6·40	4·7100

Case 1 588

H (km)	α (km/s)	β (km/s)	ρ (g/cm^3)
4·1	1·520	0·000	1·030
0·2	2·100	0·500	2·100
1·0	5·070	2·930	2·670
4·0	6·700	3·870	3·000
3·0	8·200	4·730	3·340
33·0	7·790	4·500	3·340
173·0	8·170	4·300	3·442
100·0	8·490	4·600	3·526
90·0	8·810	4·800	3·604
90·0	9·320	5·192	3·765
100·0	9·970	5·492	4·010
100·0	10·480	5·790	4·230
100·0	10·850	6·030	4·410
100·0	11·120	6·200	4·545
100·0	11·330	6·315	4·640
100·0	11·490	6·400	4·710
100·0	11·640	6·465	4·770
100·0	11·780	6·531	4·827
	11·920	6·591	4·882

kH	T (s)	C (km/s)	U (km/s)
0·02056	256·023	4·8936	3·72
0·02165	246·159	4·8336	3·68
0·02282	236·449	4·7736	3·65
0·02410	226·810	4·7136	3·63
0·02549	217·162	4·6536	3·61
0·02704	207·426	4·5936	3·60
0·02877	197·521	4·5336	3·60
0·03081	186·898	4·4736	3·59
0·03301	176·829	4·4136	3·61
0·03568	165·819	4·3536	3·62
0·03892	154·169	4·2936	3·64
0·04295	141·668	4·2336	3·67
0·04830	127·789	4·1736	3·70
0·05558	112·668	4·1136	3·74
0·06723	94·531	4·0536	3·79
0·09243	69·785	3·9936	3·87
0·11699	55·439	3·9718	3·91
0·17266	37·679	3·9596	3·95
0·22748	28·626	3·9558	3·94
0·26124	24·952	3·9520	3·92
0·31323	20·866	3·9414	3·86
0·38840	16·959	3·9110	3·70
0·43398	15·297	3·8806	3·54
0·46753	14·311	3·8502	3·39
0·51643	13·164	3·7894	3·05
0·53558	12·796	3·7590	2·88
0·55254	12·504	3·7286	2·71

velocity of 4·5 km/s requires a significant change in Poisson's ratio from the conventional value of 0·25 to a value of 0·28.

Let us now examine the possibility of a theoretical model which is compatible both with the 8·2 km/s compressional arrival in the upper mantle and the normal value of 0·25 for Poisson's ratio. In Case 1 588 the average shear velocity in the

upper mantle is still the required 4·5 km/s except that a 3 km layer of compressional velocity 8·2 km/s and shear velocity 4·73 km/s has been added just below the Mohorovičić discontinuity. This theoretical model preserves the value of 0·25 for Poisson's ratio in the upper mantle and the 3 km layer is of sufficient thickness to give the required refraction arrival yet not influence the dispersion of Rayleigh waves. Case 1 588 gives identical dispersion values with Case 1 103 and either case can be considered an excellent fit to the observed dispersion data. Whether the shear arrival from the mantle (*Sn*) is actually 4·73 km/s or 4·5 km/s cannot actually be said at this time but since *Pn* is fixed at 8·2 km/s the latter hypothesis requires a significant difference in Poisson's ratio between continental and oceanic mantle. It is therefore most important to determine by more refined explosion techniques whether the 4·73 km/s layer is actually present as this cannot be determined from surface wave data.

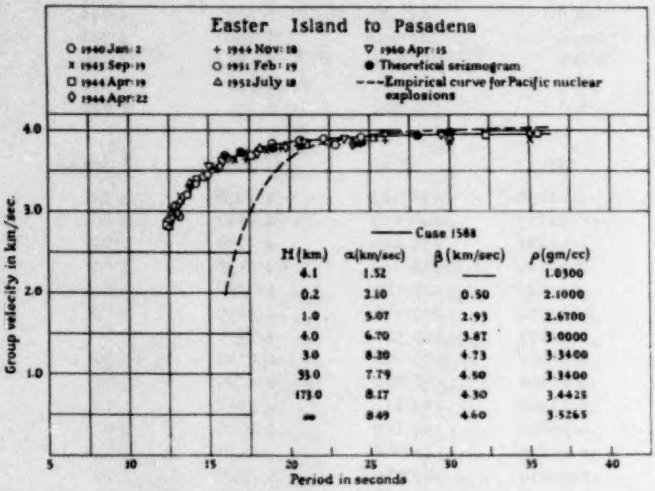


FIG. 3.—Experimental Rayleigh wave data, Easter Island to Pasadena, compared to nuclear explosion and theoretical seismogram data.

Figure 3 shows the observed Easter Island data contrasted with an empirical curve based on Pacific nuclear explosions and recorded at Pasadena. The nuclear explosion data (Griggs & Press 1959) are from an "average" Pacific path and, in addition, represent clean data since there is no origin time or epicentre location uncertainty so that a valid comparison can be made. As can be seen from Figure 3 the explosion data deviate from the observed data beginning at a period of about 25 s. (For periods less than 25 s the empirical curve differs markedly from the observed data, but this is due solely to the difference in water depth between the two paths compared.) Primarily, it should be noted that the explosion data show a group velocity maximum of about 4·05 km/s, whereas the Easter Island group velocity data show a maximum of about 3·92 km/s. Since the group velocity maximum from the explosion data is in agreement with the group velocity maximum for Dorman's Case 8 099 for the Pacific Ocean and not with the maximum shown by the Easter Island data, a conclusion can be made that there is a regional difference in the crust–mantle structure between the two Pacific Ocean paths

discussed here. This difference is interpreted as a difference in structure of the low-velocity zone in the mantle. The Easter Island to Pasadena data favour a shallower low-velocity zone than is indicated by the explosion data and Dorman's Case 8099. Furthermore, the mean shear velocity in the upper mantle for the Easter Island path is 4·5 km/s rather than 4·6 km/s indicated by the explosion data and the previous study of Dorman, Ewing & Oliver (1960) for other parts of the Pacific Ocean.

Dorman, Ewing & Oliver (1960) in their earlier study discussed the observed minimum group velocity of oceanic mantle Rayleigh waves at a period of about 225 s. Even though the data presented here are not in this period range, a check was made whether our Case 1588 would still be in agreement with Dorman's Case 8099, the best available model for oceanic mantle waves, for periods of about 200 s. Case 1588 was computed for 18 layers, as shown in Table 3, using an IBM 704 program written by Mr C. Seafeldt of the Jet Propulsion Laboratory. The tabulation of the results are also presented in Table 3 and Case 1588 is in excellent agreement with Dorman's Case 8099 from a period of about 100 s to about 260 s.

In reading seismograms by the conventional graphical method the possibility of errors exists which it is desirable to check. These errors may be caused by the group time delay of the seismograph, the approximate nature of the period–travel time determination, and the presence of Airy phases and oscillations due to the Airy Integral (Sutton, Ewing & Major 1960, Yanovskaya 1959). It is therefore necessary to check the accuracy of our dispersion data. The procedure used is to construct a theoretical seismogram utilizing a digital computer programme developed by Aki (1960). Aki's programme computes the integral

$$g(t) = 2\int_{\omega_1}^{\omega_2} \cos[\omega t - \delta(\omega)]\, d\omega$$

where

$$\begin{aligned} \delta(\omega) &= \omega\Delta/C(\omega) + \delta_{in}(\omega) \\ g(t) &= \text{impulse response} \\ \omega &= \text{angular frequency} \\ \Delta &= \text{travel distance} \\ C(\omega) &= \text{phase velocity} \\ \delta_{in}(\omega) &= \text{instrumental phase delay} \end{aligned}$$

on the Bendix G-15D digital computer. The response of the seismograph, the epicentral distance, and the theoretical phase velocity curve for the structure in Figure 3 were entered as initial data and then the programme computed the impulse response for the model selected and a theoretical seismogram, automatically plotted by the computer, was obtained. This is shown in Figure 4. The features to be noted are the large amplitude of the Airy phase, and the presence of mantle Rayleigh waves which are present as a low frequency "rider" on the fundamental wave train. The theoretical seismogram was then read in the conventional manner to obtain the data points shown in black in Figure 3. These data agree with the observed data within the precision of the observations made and thus give increased confidence in the conclusions made for the Easter Island to Pasadena path.

A number of Easter Island shocks (Table 4) were also recorded at Huancayo, Peru (Figure 1). Since these shocks were recorded on a short-period Benioff

vertical component seismograph ($T_0 = 0{\cdot}5$ s, $T_g = 4{\cdot}3$ s) no Rayleigh waves were measured with periods greater than about 27 s. Approximately 7 per cent of the great circle path from the Easter Island area to Huancayo is continental so a correction was applied for this portion of the path. Because of the uncertainty of the crustal structure of the Andes and the shortness of the great circle path, the

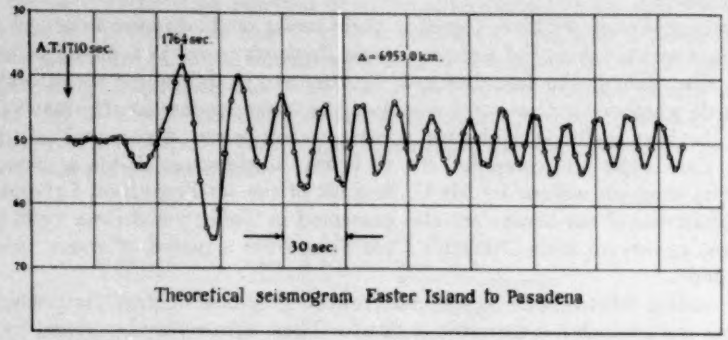


FIG. 4.—Theoretical seismogram, Easter Island to Pasadena, as plotted automatically by Bendix G-15D digital computer.

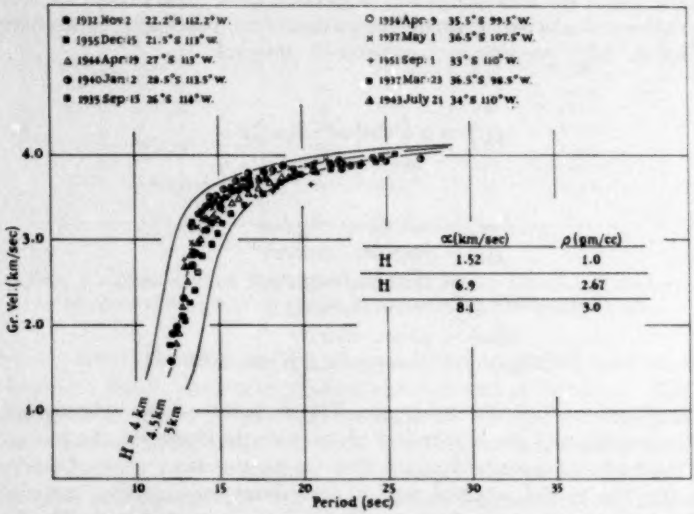


FIG. 5.—Rayleigh wave dispersion data Easter Island to Huancayo, Peru.

corrected data may be in error by as much as 0·1 km/s and no detailed interpretation of crustal and mantle structure can be made. The resulting data are shown in Figure 5 together with theoretical curves computed by Jardetzky & Press (1953) which are adequate to explain the data. Since no definite statement can be made as to the structure along the Easter Island to Huancayo path, the results are interpreted only to yield mean unconsolidated sediment thicknesses (Table 4).

In a region with mean water depth of 3·9 km/s the average unconsolidated sediment thickness is 0·57 km.

Table 4

Easter Island shocks recorded at Huancayo, Peru

Date	Time (G.C.T.)			Epicentre		Distance in km to Huancayo	Average water depth	Mean sediment thickness
	h	min	s	S	W		km	km
1932 November 2	11	03	30	22·2°	112·2°	4069	4·1	0·4
1934 April 9	15	29	29	35·5	99·5	3558	3·7	0·6
1935 September 15	14	09	08	26	114	4333	4·0	0·4
1937 March 23	00	44	21	36·5	98·5	3567	3·7	0·6
1937 May 1	12	20	53	36·5	107·5	4206	3·8	0·6
1940 January 2	11	07	18	28·5	113·5	4357	3·7	0·6
1943 July 21	04	13	57	34	110	4222	3·7	0·7
1944 April 19	22	32	08	27	113	4264	4·0	0·6
1951 September 1	08	49	18	33	110	4190	3·9	0·6
1955 December 15	01	19	02	21·5	113	4138	4·1	0·6

Figure 1 also shows the location of the Indian Ocean and Coral Sea shocks recorded at Wilkes, Antarctica. Listed in Table 5 are the origin times and epicentral distances for these earthquakes. Shocks were selected which only had a very long oceanic path to Wilkes so that errors in epicentre location and origin time would be minimized. Furthermore, since only about 1½ per cent of the great circle paths for these shocks are continental any correction applied to the observed dispersion is also small. The recording apparatus at Wilkes consisted of a three-component Press–Ewing seismograph with a seismometer free period of 15 s and a galvanometer free period of about 90 s. Because of severe operating conditions the free period of the galvanometer was later lowered to about 30 s.

Table 5

Indian Ocean and Coral Sea earthquakes

Date	Epicentre		Origin time (G.C.T.)			Distance to Wilkes
			h	min	s	km
1957 June 26	7½° S	85½° E	02	47	36	6807
1957 November 2	13 S	166½ E	18	30	24	7212
1957 November 28	15 S	168½ E	20	50	10	7086
1958 January 13	11½ N	92½ E	20	14	27	8764
1958 September 5	5 S	102 E	13	01	55	6840
1958 November 22	10½ S	112½ E	00	04	20	6201

Figure 6 shows the corrected observed Rayleigh wave dispersion data for the Indian Ocean and Coral Sea shocks. The correction to the observed dispersion data for the small continental portion of the great circle paths did not change the observed data by more than 0·02 km/s at any specific period. Also shown in

Figure 6 are theoretical curves computed for Case 1588 for different thicknesses of the water layer, all of the solid layers remaining constant. For periods greater than about 25s the data are very consistent and agree well with the theoretical curve for Case 1588.

Some limited seismic refraction work has been done in the Indian Ocean as shown in Figure 1 (Gaskell, Hill & Swallow 1958). However, since the seismic refraction lines were limited to a length of 1½ miles no depths to the Mohorovičić discontinuity were determined. Gaskell's data are not reproduced here but the deep water stations showed a crustal structure similar to that found in the Pacific Ocean.

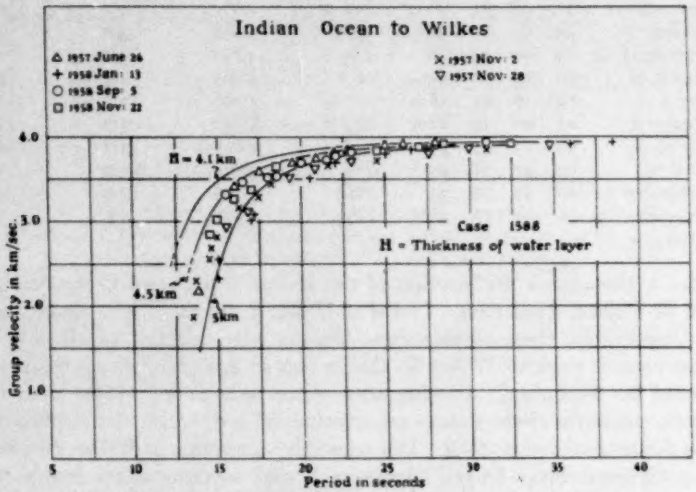


FIG. 6.—Indian Ocean Rayleigh wave dispersion data with theoretical dispersion curves for Case 1588.

The observed Rayleigh wave dispersion data in the period range from about 38s to 25s agree well with the theoretical model selected for the Easter Island to Pasadena path. On the basis of the data presented here it is indicated that the Indian Ocean has an upper mantle structure somewhat different from that found for normal areas of the Pacific Ocean. Aki (personal communication) in a study of mantle Rayleigh waves has arrived at a similar conclusion that the mantle under the Indian Ocean differs from that under the Pacific Ocean.

The Easter Island and Indian Ocean data agree in maximum observed group velocity only. The precision of the Easter Island data is good enough to require the reduction of the mean shear velocity in the upper mantle to 4·5 km/s. However, the Indian Ocean data are not good enough for us to take a position on this shear velocity reduction and the Easter Island and Indian Ocean data agree only in their need for a shallower low-velocity zone than is indicated for the Pacific Ocean basin. Nevertheless, it still must be stated that a shallower low-velocity zone is not a unique modification, but it does give a crustal model compatible with the surface wave and the seismic refraction data.

As was stated earlier, two theoretical models (Case 1103 and Case 1588) adequately explain the Easter Island to Pasadena dispersion data. Both theoretical

models required that the mean shear velocity in the upper mantle be lowered to 4·5 km/s. This low value of shear velocity implies a "soft"* mantle along the Easter Island to Pasadena path. Since the mean water depth along this path is shallower than that for the deep ocean basins, this is an indication of elevated ocean bottom topography which may be suggestive of a special structure along the Easter Island to Pasadena path. The East Pacific Rise, a topographic feature in the Pacific Ocean similar to the Mid-Atlantic Ridge, roughly trends north from the Easter Island area and Heezen (1960) believes that this rise continues to California. It may well be that the wave path considered here reflects this structural trend. Also associated with the East Pacific Rise is a relatively narrow band of high heat flow for about 6 000 km along its length (Von Herzen 1959).

Characteristic of the Mid-Atlantic Ridge are low compressional velocities measured from the mantle. At two of Raitt's refraction stations on the East Pacific Rise normal mantle velocities of 8·2 km/s were also not measured and there is the possibility that the velocity of 8·2 km/s is actually never achieved along the Easter Island to Pasadena great circle path. Nevertheless, the interesting possibility presents itself that a low mean shear velocity in the upper mantle may reflect areas of elevated temperatures or magma chambers which may be involved with the volcanics of oceanic ridge structures.

Acknowledgments

This research was partially supported by National Science Foundation I.G.Y. grants 5 767 and 8 994.

Seismological Laboratory,
California Institute of Technology,
Pasadena,
Cal., U.S.A.
1960 *September.*

References

Aki, K., 1960. Study of earthquake mechanism by a method of phase equalization applied to Rayleigh and Love waves. *J. Geophys. Res.*, **65,** 729–740.

Dorman, J., Ewing, M. & Oliver, J., 1960. Study of the shear-velocity distribution in the upper mantle by mantle Rayleigh waves. *Bull. Seismol. Soc. Amer.*, **50,** 87–115.

Gaskell, T. F., Hill, M. N. & Swallow, J. C., 1958. Seismic measurements made by H.M.S. *Challenger* in the Atlantic, Pacific, and Indian Oceans and in the Mediterranean Sea, 1950–53. *Phil. Trans. Roy. Soc A*, **251,** 23–83.

Griggs, D. T. & Press, F., 1959. Probing the Earth with nuclear explosions. *Research Memorandum* 2456 *AEC*, 1–38.

Heezen, B., 1960. Geologie Sousmarine et displacement continental. 83rd Colloque, Centre Nat. Res. Sci., Paris.

* "Soft" is defined here to signify a higher ratio of α/β or a mean lower value of β than that indicated for the rest of the Pacific Ocean basin.

Jardetzky, W. S. & Press, F., 1953. Crustal structure and surface wave dispersion. Part 3: Theoretical dispersion curves for suboceanic Rayleigh waves. *Bull. Seismol. Soc. Amer.*, **43**, 137–144.

Oliver, J., Ewing, M., & Press, F., 1955. Crustal structure and surface wave dispersion. Part 4: Atlantic and Pacific Ocean basins. *Bull. Geol. Soc. Amer.*, **66**, 913–946.

Oliver, J., Kovach, R. & Dorman, J., 1960. Crustal structure of the New York–Pennsylvania area (in preparation).

Press, F., 1960. Crustal structure in the California–Nevada region. *J. Geophys. Res.*, **65**, 1039–1051.

Raitt, R. W., 1956. Seismic refraction studies of the Pacific Ocean basin. *Bull. Geol. Soc. Amer.*, **67**, 1623–1640.

Sutton, G. H., Ewing, M. & Major, M., 1960. Rayleigh wave group velocity extrema. Paper presented orally at the International Association of Seismology and Physics of the Earth's Interior, Helsinki.

Von Herzen, R., 1959. Heat flow values from the southeastern Pacific. *Nature*, **183**, 882–883.

Yanovskaya, T. B., 1959. An investigation of dispersing surface waves in the region of minimum group velocity. *Izvest. Akad. Nauk. SSSR (Geophys. Ser.)*, **12**, 1753–1760.

IV. Free Oscillations

Z. Alterman, H. Jarosch & C. L. Pekeris, Propagation of Rayleigh waves in the Earth 219

N. Jobert, Calcul approché de la période des oscillations sphéroïdales de la terre 242

H. Takeuchi, Torsional oscillations of the Earth 259

Propagation of Rayleigh Waves in the Earth

Z. Alterman, H. Jarosch and C. L. Pekeris

"As when the massy substance of the Earth quivers."—*C. Marlowe, Tamburlaine the Great.* Pt. 2, I. 1.

Summary

The propagation of Rayleigh waves in the Earth is investigated in the whole range of periods T from about 10 s up to one hour. Three methods are necessary in order to cover this range of periods effectively. The standard *flat Earth method*, with neglect of gravity, gives values for the phase velocity C correct to within 1 per cent up to $T = 50$ s only, and for the group velocity U up to $T = 250$ s. The method of the *flattening of the Earth*, with neglect of gravity, has the 1 per cent accuracy limits for C and U at 300 and 400, respectively. Inclusion of gravity effects in the flattening of the Earth approximation does not alter the above limits. For $T > 300$ ($n < 25$) one must determine the period $T(n)$ of free oscillation of the Earth as a function of the order of the spherical harmonic n. This involves the solution of a system of differential equations of the sixth order, in which the gravitational effects are included. The wave penetrates appreciably into the core already at $T = 600$. Using the above three methods in their respective ranges of validity, we have evaluated $C(T)$ and $U(T)$ for (1) Bullen's Model B, (2) the Jeffreys–Bullen Model, as modified by Dorman, Ewing and Oliver, and (3) the Gutenberg Model. The observed Rayleigh group velocity data of Ewing and Press for $T < 380$ s and the phase velocity data of Nafe and Brune for $T < 300$ s agree with the values computed for the Gutenberg model, but not for the other models. This substantiates a previous conclusion reached by Takeuchi, Press and Kobayashi and by Dorman, Ewing and Oliver that the observed Rayleigh wave data provide evidence in support of Gutenberg's low-velocity layer. The few observed Rayleigh group velocities between $T = 400$ and 600 are substantially lower than the theoretical values for all the three models.

1. Introduction

In a previous investigation (Alterman, Jarosch & Pekeris 1959) subsequently referred to as I, methods were developed for determining the periods of free oscillation of the Earth as represented by models based on earthquake data and on other geophysical information concerning the internal constitution of the Earth. Our initial aim was to test a conjecture of Benioff (1954) that the 57-min

oscillation which he observed on the records of the Kamchatka earthquake of 1952 is a free spheroidal oscillation of the Earth. This study, which dealt primarily with oscillations of spherical harmonic order $n = 2$, was then extended also to the dynamical theory of Earth tides (I, and Pekeris & others 1959). Recent advances (Benioff & Press 1958) made in the design of long-period seismographs hold out the hope of extending the observed seismic spectrum up to the one-hour limit. Already Ewing and Press (Benioff & Press 1958, and Ewing & Press 1956) have succeeded in recording Rayleigh waves of a period T of 630s. The propagation of waves of periods of 600s and higher is governed by the free modes of oscillation of the Earth of spherical harmonic order n ranging from 10 to 2. The amplitude of these modes extends from the surface of the Earth down into the core and their motion is governed not only by the elastic restoring forces but also by forces arising from the perturbed gravitational field. The analysis of these modes has therefore to be made by solving the systems of 6 differential equations I(28) to I(33) and I(35) to I(39) in the mantle and core respectively. This *complete normal mode method* must be used for periods $T > 400$s, ($n < 17$). In the range of $400 > T > 300$ ($17 < n < 25$) the amplitude of the normal modes becomes negligibly small throughout the core, so that it is sufficient to carry out the integrations only from the bottom of the mantle to the surface. In this *restricted normal mode method* one imposes the condition at the core boundary of the vanishing of all of the components of the displacements, and of the stresses.

In the period-range of $300 > T > 50$ ($25 < n < 200$) the analysis can be considerably simplified by taking account of the approach to flat Earth conditions, without, however, altogether neglecting the still appreciable effect of the sphericity of the Earth. This is accomplished by the use of the *Earth-flattening approximation*. In this method the space metric is transformed so as to make the surface of the Earth flat and to curve rays which were originally straight. The effect of the sphericity of the Earth turns out to be equivalent to the superposition on the given velocity–depth function of a perturbation term which varies linearly with depth. In the period ranges where the Earth-flattening approximation applies we have found that the gravitational forces may be neglected.

Finally, for periods $T < 50$s, the standard flat Earth approximation with neglect of gravity is valid. This method has been used extensively by Stoneley (1953), and recently by Press & Takeuchi (1960), and by Dorman, Ewing & Oliver (1960). The regions of applicability of the various methods are summarized in Table 1.

Table 1

Ranges of applicability of various methods

Range of period T in seconds	Range of n	Method	Core included	Gravity included
$3220 > T > 400$	$17 > n > 2$	Complete normal mode	Yes	Yes
$400 > T > 300$	$25 > n > 17$	Restricted normal mode	No	Yes
$300 > T > 50$	$200 > n > 25$	Earth flattening approximation	No	No
$50 > T$	$n > 200$	Flat earth	No	No

The limits set for the applicability of the various approximate methods were obtained by comparing results obtained for the phase velocity C with those derived from the complete normal mode theory. It turns out that the group velocities U come out more accurately by the approximate methods than do the phase velocities C.

We have carried out the analysis of propagation of Rayleigh waves in the Earth for three models whose properties are described in Tables 2, 3, and 4 and in Figures 1 and 2. The "Jeffreys–Bullen" Model is taken from a paper by Dorman, Ewing & Oliver (1960). In their method of calculation of dispersion of Rayleigh waves, the Earth is represented by a system of homogeneous layers. They have determined the properties of these layers so as to follow the average properties of the Jeffreys–Bullen Model. The "Bullen" Model is based on the data given (Bullen 1950) for Bullen's Model B, with linear interpolation at intermediate points. The properties of the top layer of 33 km thickness are taken from the paper by Press & Takeuchi (1960). The data on the step-function model designated as the "Gutenberg" Model are taken from the paper by Dorman, Ewing & Oliver (1960) and from Bullen's book *An Introduction to the Theory of Seismology*, p. 218. This model differs from the other two by having a *low-velocity layer*.

2. The normal mode solution

In studying the propagation of Rayleigh waves of periods $T > 300$ s we must follow the method of normal modes described in I. The periods of free spheroidal oscillations for Bullen's Model B were there given in Table 4 for $n = 2$, 3, and 4. We have extended this analysis to higher values of n, and the results are shown in Table 5 of the present paper. The direct result of the eigenvalue problem is the period $T(= 2\pi/\sigma)$ for a given value of n. The normal mode is then represented by a standing oscillation of the form

$$V = \exp(i\sigma t)F(r)P_n(\cos\theta). \tag{1}$$

Such a standing oscillation can be conceived as the result of a superposition of two waves travelling in opposite directions. For large n this decomposition can be accomplished by the use of the asymptotic expansion of $P_n(\cos\theta)$:

$$V \to F(r)(2n\pi\sin\theta)^{-\frac{1}{2}}\left\{\exp\left(i\left[\sigma t-(n+\tfrac{1}{2})\theta+\frac{\pi}{4}\right]\right)+\right.$$
$$\left.+\exp\left(i\left[\sigma t+(n+\tfrac{1}{2})\theta-\frac{\pi}{4}\right]\right)+O(1/n)\right\}. \tag{2}$$

The first term in the braces represents a wave travelling in the positive θ-direction:

$$\sigma t-(n+\tfrac{1}{2})\theta = \sigma t-\frac{x}{C}, \tag{3}$$

$$x = a\theta, \qquad C = \frac{a\sigma}{(n+\tfrac{1}{2})}, \qquad k = \frac{\sigma}{C} = \frac{(n+\tfrac{1}{2})}{a}. \tag{4}$$

A group velocity U can be derived from

$$U = \frac{d(kC)}{dk} = a\frac{d\sigma}{dn}. \tag{5}$$

The values of U given in Table 5 were obtained by finite differencing of $\sigma(n)$. For

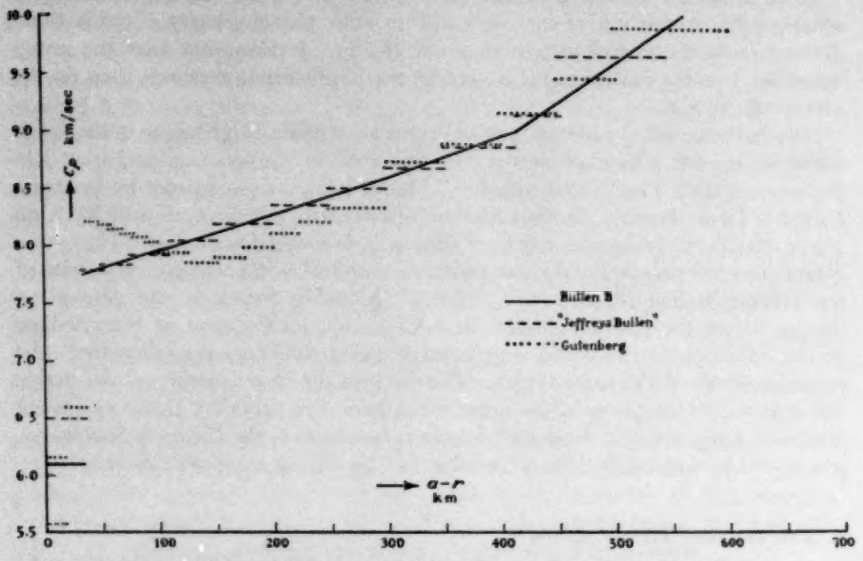


FIG. 1.—Variation with depth $a-r$ of the compressional velocity C_p for models Bullen B, "Jeffreys–Bullen" and Gutenberg.

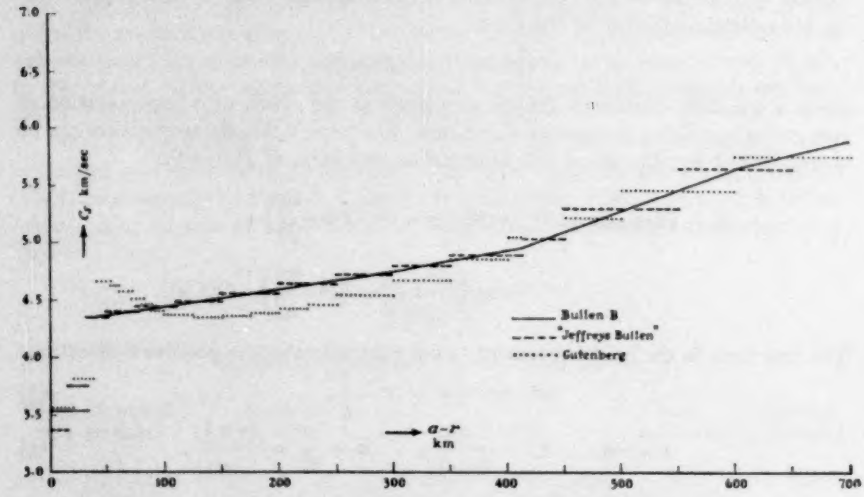


FIG. 2.—Variation with depth $a-r$ of the shear velocity C_s for models Bullen B, "Jeffreys–Bullen" and Gutenberg.

Table 2

Properties of the Bullen B Earth Model

$a-r$	ρ	C_p	C_s	μ	λ	g
km	g/cm³	km/s	km/s	10^{11} dyn/cm²	10^{11} dyn/cm²	cm/s²
0	2·76	6·10	3·54	3·45	3·37	982
33	2·82	6·10	3·54	3·52	3·45	985
33	3·32	7·75	4·35	6·28	7·38	985
80	3·36	7·89	4·42	6·56	7·77	986
80	3·87	7·89	4·42	7·56	8·95	986
100		7·95	4·45			
200	3·94	8·26	4·60	8·34	10·21	985
300		8·58	4·76			
400	4·06	8·93	4·94	9·90	12·55	983
413		8·97	4·96			
600	4·18	10·25	5·66	13·39	17·13	980
800	4·30	11·00	6·13	16·16	19·71	978
1 000	4·41	11·42	6·36	17·84	21·84	976
1 200	4·52	11·71	6·50	19·10	23·79	976
1 400	4·63	11·99	6·62	20·29	25·98	976
1 600	4·74	12·26	6·73	21·47	28·31	978
1 800	4·84	12·53	6·83	22·58	30·83	982
2 000	4·94	12·79	6·93	23·72	33·36	987
2 200	5·03	13·03	7·02	24·79	35·82	997
2 400	5·13	13·27	7·12	26·01	38·32	1 010
2 600	5·22	13·50	7·21	27·14	40·86	1 029
2 700	5·27	13·57	7·26	27·74	41·57	1 042
2 800		13·64	7·30			
2 898		13·64	7·30			
2 898		8·10				
2 900	5·57			29·68	44·26	1 069
2 900	9·74				63·94	1 069
3 000	9·90	8·22			66·89	1 041
3 200	10·20	8·47			73·18	1 006
3 400	10·47	8·76			80·34	960
3 600	10·72	9·04			87·61	913
3 800	10·95	9·28			94·30	865
4 000	11·16	9·51			100·93	815
4 200	11·36	9·70			106·89	765
4 400	11·54	9·88			112·65	718
4 600	11·71	10·06			118·51	670
4 800	11·87	10·25			124·71	628
4 982	12·00	10·44			130·79	599
5 121	15·00	9·47			134·52	563
5 121	15·00	11·16			186·82	563
5 700		11·26				
6 371	17·90	11·31			228·92	0

Table 3

Properties of Jeffreys–Bullen Earth Model as modified by Dorman, Ewing & Oliver (1960)

$a-r$ km	ρ g/cm³	C_p km/s	C_s km/s	μ 10^{11} dyn/cm²	λ 10^{11} dyn/cm²
0	2·65	5·57	3·36	2·99	2·24
15	2·87	6·50	3·74	4·01	4·10
33	3·33	7·78	4·36	6·33	7·47
50	3·35	7·83	4·39	6·46	7·63
75	3·37	7·92	4·44	6·64	7·85
110	3·41	8·04	4·49	6·87	8·29
150	3·45	8·19	4·56	7·17	8·79
200	3·49	8·35	4·64	7·51	9·31
250	3·53	8·50	4·72	7·86	9·78
300	3·57	8·67	4·80	8·23	10·38
350	3·62	8·86	4·90	8·68	11·02
413	3·70	9·14	5·04	9·40	12·11
450	3·89	9·65	5·31	10·97	14·29
550	4·13	10·25	5·66	13·21	16·91
650	4·32	10·68	5·93	15·19	18·89
750	4·49	11·00	6·13	16·87	20·58
850	4·62	11·28	6·29	18·28	22·23
1 000	4·74	11·57	6·44	19·65	24·13
1 200	4·92	11·99	6·62	21·54	27·58
1 600	5·14	12·53	6·83	23·95	32·71
2 000	5·34	13·03	7·02	26·32	38·03
2 400	5·54	13·50	7·21	28·80	43·37
2 800	5·69	13·64	7·30	30·32	45·22

Table 4

Properties of the Gutenberg Earth Model

$a-r$ km	ρ g/cm³	C_p km/s	C_s km/s	μ 10^{11} dyn/cm²	λ 10^{11} dyn/cm²	g cm/s²
0	2·74	6·14	3·55	3·45	3·42	982
19	3·00	6·58	3·80	4·33	4·32	983
38	3·32	8·20	4·65	7·18	7·97	984
50	3·34	8·17	4·62	7·13	8·04	985
60	3·35	8·14	4·57	7·00	8·20	985
70	3·36	8·10	4·51	6·83	8·38	986
80	3·37	8·07	4·46	6·70	8·54	986
90	3·38	8·02	4·41	6·57	8·59	986
100	3·39	7·93	4·37	6·47	8·37	986
125	3·41	7·85	4·35	6·45	8·11	987
150	3·43	7·89	4·36	6·52	8·31	988
175	3·46	7·98	4·38	6·64	8·76	989
200	3·48	8·10	4·42	6·80	9·23	989
225	3·50	8·21	4·46	6·96	9·67	990
250	3·53	8·38	4·54	7·28	10·24	991
300	3·58	8·62	4·68	7·84	10·92	992
350	3·62	8·87	4·85	8·52	11·45	993
400	3·69	9·15	5·04	9·37	12·15	995
450	3·82	9·45	5·21	10·37	13·38	996
500	4·01	9·88	5·45	11·91	15·32	997
600	4·21	10·30	5·76	13·97	16·73	998
700	4·40	10·71	6·03	16·00	18·47	998
800	4·56	11·10	6·23	17·70	20·79	997
900	4·63	11·35	6·32	18·49	22·66	995
1 000	4·74	11·60	6·42	19·54	24·71	993
1 200	4·85	11·93	6·55	20·81	27·41	990
1 400	4·96	12·17	6·69	22·20	29·06	986

Table 4—*continued*

$a-r$ km	ρ g/cm³	C_p km/s	C_s km/s	μ 10^{11} dyn/cm²	λ 10^{11} dyn/cm²	g cm/s²
1600	4·96	12·17	6·69	22·20	29·06	986
1800	5·07	12·43	6·80	23·44	31·45	983
2000	5·19	12·67	6·90	24·71	33·90	982
2200	5·29	12·90	6·97	25·70	36·63	981
2400	5·39	13·10	7·05	26·79	38·92	984
2600	5·49	13·32	7·15	28·07	41·27	989
2800	5·59	13·59	7·23	29·22	44·80	997
	5·69	13·70	7·20	29·50	47·80	1011
2898	9·40	8·10			61·71	1037
3000	9·55	8·23			64·72	1015
3500	10·15	8·90			80·42	908
4000	10·7	9·50			96·58	800
4500	11·2	9·97			111·39	631
4982	11·5	10·44			125·34	469
5121	12·0	10·75			138·66	422
6371	12·3	11·31			157·34	0

small values of n, of course, the expansion (2) is not valid and the usual concept of group-velocity requires modification. The distribution of amplitude with depth in the normal modes is shown in Figures 3 to 6. It is seen that for $n = 10$ ($T = 591$) the amplitude in the core is still appreciable. On the other hand, for $n > 17$ the amplitude in the core becomes negligibly small, and we may confine the integration to the mantle.

3. The Earth flattening approximation

For periods around 300 s the amplitude of the free oscillations is not only negligibly small in the core but also in the lower part of the mantle. It is, however, still not permissible to neglect altogether the curvature of the Earth. Referring to Figure 7, the sphericity of the Earth makes itself felt through the circumstance that a pulse originating on the surface at the point A can reach a point C sooner by travelling along the chord ADC rather than along the arc ABC. The flattening of the Earth approximation (Pekeris 1946; Pryce 1953; Koo & Katzin 1960) transforms the problem of propagation over a sphere to one over a flat Earth while still retaining the features of the inherent geometry of the sphere. This is accomplished by transforming the space-metric so as to make the surface flat and the rays curved. If the travel time along the ray ADC in the transformed flat space of Figure 7 is to be

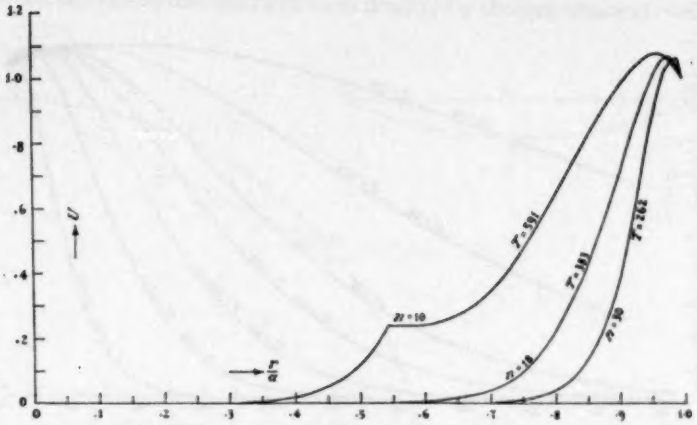


FIG. 3.—Distribution of radial displacement U in the normal spheroidal modes of orders $n = 10$, 18, and 30 in model Bullen B.

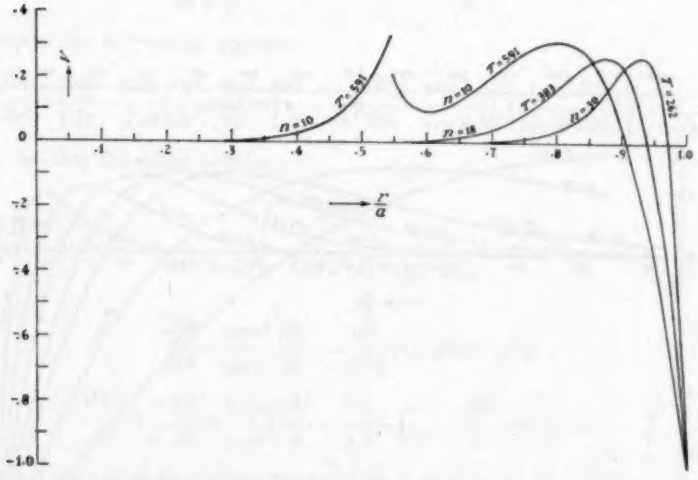


FIG. 4.—Distribution of horizontal displacement V in the normal spheroidal modes of orders $n = 10$, 18 and 30 in model Bullen B.

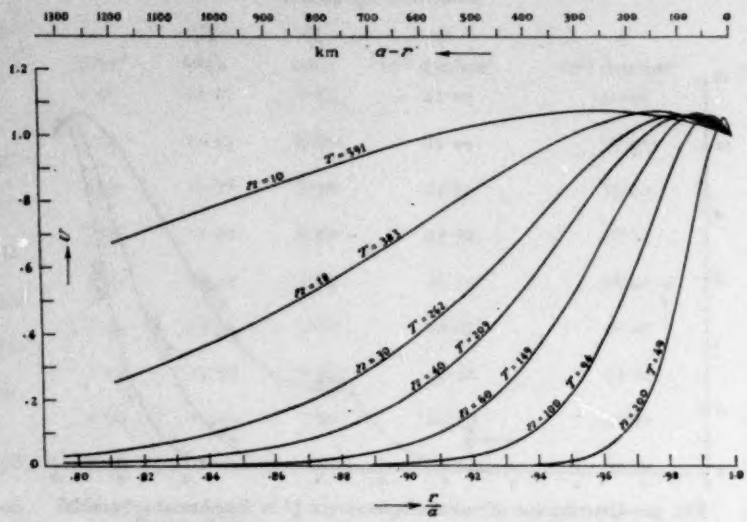


FIG. 5.—Distribution of radial displacement U in the normal spheroidal modes of orders n = 10 to 200 in model Bullen B.

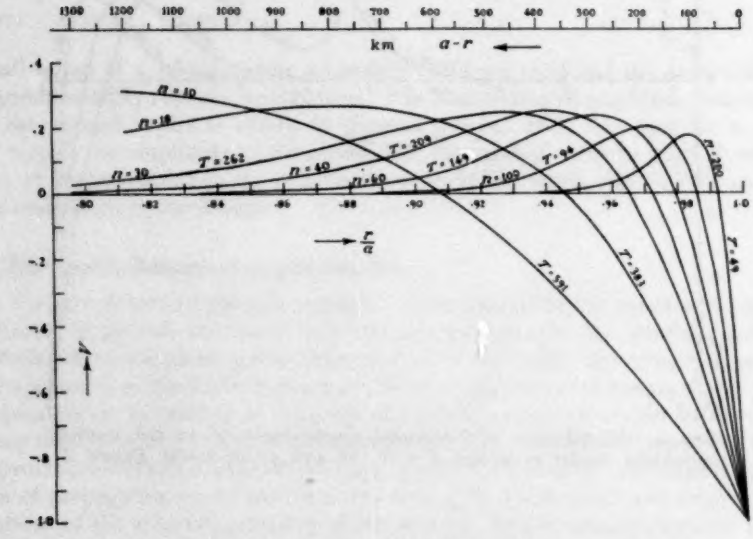


FIG. 6.—Distribution of horizontal displacement V in the normal spheroidal modes of orders n = 10 to 200 in model Bullen B.

less than along the ray ABC, then the velocity must increase with depth. The amount of this velocity variation with depth can be established from the form which the differential equations for the sphere take on for large values of n. The procedure can be demonstrated most directly for the free torsional oscillations.

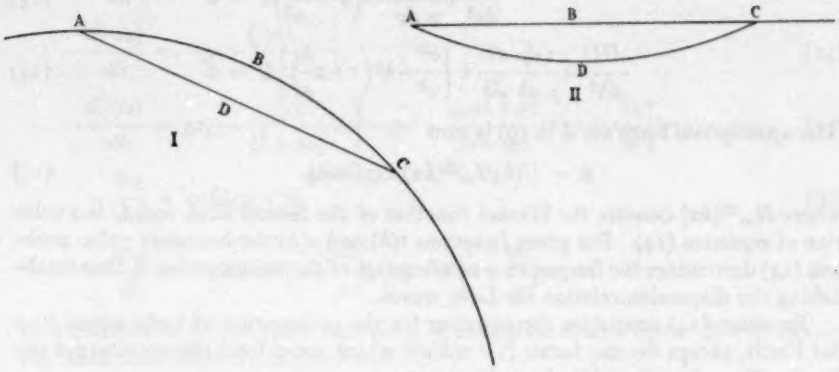


FIG. 7.—The flattening of the Earth approximation.

In that case the radial component of motion u vanishes, and the horizontal components v and w are derived from one function ψ

$$v = \frac{1}{\sin\theta}\frac{\partial\psi}{\partial\phi}\exp(i\sigma t), \qquad w = -\frac{\partial\psi}{\partial\theta}\exp(i\sigma t), \tag{6}$$

which obeys the differential equation

$$\frac{\partial^2\psi}{\partial r^2}+\frac{2}{r}\frac{\partial\psi}{\partial r}+\frac{1}{\mu}\frac{d\mu}{dr}\left(\frac{\partial\psi}{\partial r}-\frac{\psi}{r}\right)+\left[\frac{\sigma^2}{c^2}-\frac{1}{r^2}\left(\frac{\partial^2\psi}{\partial\theta^2}+\frac{\cos\theta}{\sin\theta}\frac{\partial\psi}{\partial\theta}+\frac{1}{\sin^2\theta}\frac{\partial^2\psi}{\partial\phi^2}\right)\right]=0, \tag{7}$$

where c denotes the shear velocity:

$$c^2 = \mu/\rho_0. \tag{8}$$

Putting now

$$\psi(r,\theta,\phi) = U(r)P(\theta)\exp(im\phi), \tag{9}$$

we get

$$\frac{d^2P}{d\theta^2}+\frac{\cos\theta}{\sin\theta}\frac{dP}{d\theta}-\frac{m^2}{\sin^2\theta}P+a^2k^2P=0, \tag{10}$$

$$\frac{d^2U}{dr^2}+\frac{2}{r}\frac{dU}{dr}+\frac{1}{\mu}\frac{d\mu}{dr}\left(\frac{dU}{dr}-\frac{U}{r}\right)+\left(\frac{\sigma^2}{c^2}-\frac{a^2k^2}{r^2}\right)U=0. \tag{11}$$

Here a denotes the radius of the Earth, and we have written a^2k^2 instead of the customary separation constant $n(n+1)$ with the intention of relinquishing the requirement that P be finite at $\theta = 0$ and $\theta = \pi$. The plan is to build up a solution satisfying the source-condition and the outgoing wave condition from a series which converges rapidly everywhere except near the points $\theta = 0$, and $\theta = \pi$, where each term of the series becomes singular. The limiting form of equations

(10) and (11) as $ak \to \infty$ is obtained by using as horizontal and vertical coordinates

$$x = a\theta, \qquad h = a - r, \tag{12}$$

and neglecting terms of order $1/a$, except when they are multiplied by the large factor k^2:

$$\frac{d^2P}{dx^2} + \frac{1}{x}\frac{dP}{dx} + \left(k^2 - \frac{m^2}{x^2}\right)P = 0 \tag{13}$$

$$\frac{d^2U}{dh^2} + \frac{1}{\mu}\frac{d\mu}{dh}\frac{dU}{dh} + \left[\frac{\sigma^2}{c^2} - k^2\left(1 + 2\frac{h}{a}\right)\right]U = 0. \tag{14}$$

The appropriate form for ψ in (9) is now

$$\psi = U(h)H_m^{(2)}(kx)\exp(im\phi), \tag{15}$$

where $H_m^{(2)}(kx)$ denotes the Hankel function of the second kind, and U is a solution of equation (14). For given functions $c(h)$ and $\mu(h)$ the boundary value problem (14) determines the frequency σ as a function of the wave number k, thus establishing the dispersion relation for Love waves.

Equation (14) resembles the equation for the propagation of Love waves on a flat Earth, except for the factor $[1+2(h/a)]$ which arises from the curvature of the Earth. Now since $k \simeq (\sigma/c_0)$, we have

$$\frac{\sigma^2}{c^2} - 2k^2\frac{h}{a} \simeq \frac{\sigma^2}{\{c[1+(h/a)]\}^2}. \tag{16}$$

The flattening of the Earth approximation thus effectively introduces a linear increase of velocity with depth.

In the case of spheroidal oscillations, the modifications of equations I(28) to I(33), which are introduced by the Earth flattening approximation, can be derived simply by writing

$$u = \frac{\partial\psi}{\partial r} + \frac{1}{r^2\sin\theta}\frac{\partial\Omega}{\partial\theta}, \qquad v = \frac{1}{r}\frac{\partial\psi}{\partial\theta} - \frac{1}{r\sin\theta}\frac{\partial\Omega}{\partial r}, \qquad w = 0, \tag{17}$$

$$\psi = F(r)P_n(\theta), \qquad \Omega = H(r)\sin\theta\frac{\partial P_n}{\partial\theta}. \tag{18}$$

It is then found that the functions $U(r)$ and $V(r)$ of equations I(26) and I(27) are given by

$$U(r) = \frac{dF}{dr} - \frac{n(n+1)}{r^2}H, \tag{19}$$

$$V(r) = \frac{1}{r}\left(F - \frac{dH}{dr}\right). \tag{20}$$

When (19) and (20) are substituted into equations I(28) to I(33) the flattening of the Earth approximation is effected by writing

$$\frac{n(n+1)}{r^2} \to k^2\left(1 + \frac{2h}{a}\right). \tag{21}$$

The same result can, however, be obtained directly from I(28) to I(33) by using as variables ry_3 and ry_4 instead of y_3 and y_4, as is suggested by the $1/r$ factor in (20).

We thus get

$$-\frac{dy_1}{dh} = \frac{y_2}{(\lambda+2\mu)} + \frac{\lambda}{(\lambda+2\mu)}k^2\left(1+\frac{2h}{a}\right)(ry_3) \tag{22}$$

$$-\frac{dy_2}{dh} = -\sigma^2\rho y_1 + k^2\left(1+2\frac{h}{a}\right)[\rho g(ry_3)+(ry_4)] - \rho y_4, \tag{23}$$

$$-\frac{d(ry_3)}{dh} = -y_1 + \frac{(ry_4)}{\mu}, \tag{24}$$

$$-\frac{d(ry_4)}{dh} = g\rho y_1 - \frac{\lambda y_2}{(\lambda+2\mu)} + \left[-\rho\sigma^2 + \frac{4\mu(\lambda+\mu)}{(\lambda+2\mu)}k^2\left(1+2\frac{h}{a}\right)\right](ry_3) - \rho y_5, \tag{25}$$

$$-\frac{dy_5}{dh} = 4\pi G\rho y_1 + y_6, \tag{26}$$

$$-\frac{dy_6}{dh} = -4\pi G\rho k^2\left(1+\frac{2h}{a}\right)(ry_3) + k^2\left(1+\frac{2h}{a}\right)y_5. \tag{27}$$

It is convenient to put these equations in non-dimensional form. Let λ^* denote the maximum value of λ in the Earth,

$$\lambda = \lambda^*\lambda_1 \qquad \mu = \lambda^*\mu_1, \qquad g = g(a)g_1, \tag{28}$$

$$\rho = \bar{\rho}\rho_1, \qquad b = \frac{\bar{\rho}\sigma^2a^2}{\lambda^*}, \qquad c = \frac{\bar{\rho}^2a^2C}{\lambda^*}, \tag{29}$$

$$y_1 = az_1, \qquad y_2 = \lambda^*z_2 \qquad ry_3 = a^2z_3, \tag{30}$$

$$ry_4 = a\lambda^*z_4, \qquad y_5 = a^2C\bar{\rho}z_5, \qquad y = aC\bar{\rho}z_6. \tag{31}$$

$$n(n+1) = a^2k^2 \qquad s = r/a.$$

Equations (22) to (27) then become

$$\frac{dz_1}{ds} = \frac{z_2}{(\lambda_1+2\mu_1)} + \frac{\lambda_1}{(\lambda_1+2\mu_1)}n(n+1)\left(1+\frac{2h}{a}\right)z_3, \tag{33}$$

$$\frac{dz_2}{ds} = -b\rho_1z_1 + n(n+1)\left(1+\frac{2h}{a}\right)\left[\frac{4\pi}{3}c\rho_1g_1z_3+z_4\right] - c\rho_1z_6, \tag{34}$$

$$\frac{dz_3}{ds} = -z_1 + \frac{z_4}{\mu_1}, \tag{35}$$

$$\frac{dz_4}{ds} = \frac{4\pi}{3}cg_1\rho_1z_1 - \frac{\lambda_1}{(\lambda_1+2\mu_1)}z_2 - b\rho_1z_3 + \frac{4\mu_1(\lambda_1+\mu_1)}{(\lambda_1+2\mu_1)}n(n+1)\left(1+2\frac{h}{a}\right)z_3 - c\rho_1z_5, \tag{36}$$

$$\frac{dz_5}{ds} = 4\pi\rho_1z_1 + z_6, \tag{37}$$

$$\frac{dz_6}{ds} = -4\pi n(n+1)\left(1+\frac{2h}{a}\right)\rho_1z_3 + n(n+1)\left(1+\frac{2h}{a}\right)z_5. \tag{38}$$

The boundary condition at the surface of the Earth $r = a$ is

$$z_2 = 0, \qquad z_4 = 0, \qquad z_6+(n+1)z_5 = 0. \tag{39}$$

At the boundary of the core we require the vanishing of all the z_i.

4. Discussion of results

Our results for the dispersion of Rayleigh waves in the Earth are given in Tables 5, 6 and 7 and in Figures 8 and 9 for Bullen's Model B; in Table 8 and Figure 10 for the "Jeffreys–Bullen" Model; and in Tables 9, 10 and 11 and Figure 11 for the Gutenberg Model. The regions of applicability of the various methods of approximation can be judged from the results for Bullen's Model B shown in Figure 8. The phase velocity curve C (dashed) and the group velocity curve U of the Earth flattening approximation begin to deviate from the exact normal mode curve (solid) at about $T = 250$s, while the plane Earth curve (dotted) for C is in error already beyond $T = 40$s. The group velocity curve U for the plane Earth approximation turns out to be accurate up to about 250s. Inclusion

Table 5

Period T, phase velocity C and group velocity U for Bullen's Model B obtained by the normal mode solution. The interpolation is based on the exact values, computed at intervals of 5. n denotes the order of spherical harmonic. For $n < 17$ the solution extends through the mantle and core; for $n > 17$ the amplitude in the core is negligible.

	T min		C	U		T min		C	U
n	Exact	Interpol.	km/s	km/s	n	Exact	Interpol.	km/s	km/s
2	53·70		4·97		32		4·16	4·94	
3	35·50		5·37	6·75	33		4·06	4·91	
4	25·73		5·76	7·41	34		3·97	4·88	
5	19·85		6·11	7·73	35	3·88		4·85	
6	16·12		6·37	7·64	36		3·79	4·82	
7	13·64		6·52	7·20	37		3·71	4·80	
8	11·95		6·57	6·53	38		3·63	4·77	
9	10·77		6·52	5·85	39		3·56	4·75	
10	9·88		6·43	5·33	40	3·49		4·73	
11	9·18	9·16	6·32	5·01			3·45		3·84
12	8·61	8·57	6·20	4·80	41	3·42	3·42	4·70	
13	8·11	8·08	6·09	4·66	42		3·35	4·68	
14	7·68	7·67	5·99	4·56	43		3·29	4·66	
15	7·30		5·89	4·48	44		3·23	4·65	
16	6·96	6·97	5·81	4·40	45	3·17		4·63	
17	6·66	6·66	5·72	4·33	46		3·11	4·61	
18	6·39	6·39	5·65	4·27	47		3·06	4·60	
19	6·14	6·14	5·57	4·21	48		3·00	4·58	
20	5·91		5·51	4·16	49		2·95	4·57	
21	5·70	5·69	5·44		50	2·90		4·55	3·87
22		5·48	5·41		51		2·85	4·54	
23		5·29	5·37		52		2·81	4·53	
24		5·12	5·32		53		2·76	4·51	
25	4·96		5·28		54		2·72	4·50	
26		4·82	5·22		55	2·68		4·49	3·88
27		4·70	5·16		56		2·64	4·48	
28		4·58	5·11		57		2·60	4·47	
29		4·47	5·06		58		2·56	4·46	
30	4·37		5·01		59		2·52	4·45	
		4·31		3·88	60	2·48		4·44	
31	4·26	4·26	4·98				2·47		3·89
					61	2·45	2·45	4·43	

Table 6

Earth flattening approximation for Bullen's Model B
σ—frequency, T—period, C—phase velocity, U—group velocity
$n(n+1) = a^2k^2$

n	σ	T	C	U	n	σ	T	C	U
	s^{-1}	s	km/s	km/s		s^{-1}	s	km/s	km/s
10	0·0106909	588	6·49		151	0·0981202	64	4·13	
		565		5·63	200	0·127154	49	4·04	
11	0·0115744	543	6·41				49		3·71
					201	0·127736	49	4·04	
12	0·0124334	505	6·34		250	0·155644	40	3·96	
		490		4·98			40		3·55
13	0·0132144	475	6·24		252	0·156757	40	3·96	
					300	0·182878	34	3·88	
15	0·0146324	429	6·01				34		3·39
		420		4·50	302	0·183941	34	3·87	
16	0·0153382	410	5·92		400	0·233856	27	3·72	
20	0·0180421	348	5·61				27		3·12
		342		4·16	402	0·234835	27	3·72	
21	0·0186946	336	5·54		500	0·281569	22	3·58	
25	0·0211656	297	5·29				22		2·98
		292		4·11	504	0·283439	22	3·58	
26	0·0218111	288	5·24		600	0·327989	19	3·48	
30	0·0243188	258	5·08				19		2·95
		255		3·88	604	0·329841	19	3·48	
31	0·0249272	252	5·04		806	0·424479	15	3·35	
40	0·0303591	207	4·78				15		3·03
		205		3·84	808	0·425431	15	3·35	
41	0·0309614	203	4·75		1000	0·518212	12	3·30	
60	0·0424755	148	4·47				12		3·12
		147		3·89	1003	0·519681	12	3·30	
61	0·0430858	146	4·46				12		3·12
100	0·0670392	94	4·25		1005	0·520661	12	3·30	
		93		3·92					
101	0·0676538	93	4·25						
150	0·0975178	64	4·13						
		64		3·84					

Table 7

Plane Earth approximation for Bullen Model B
σ—frequency, T—period, C—phase velocity, U—group velocity
$n(n+1) = a^2k^2$

n	σ	T	C	U
	s^{-1}	s	km/s	km/s
10	0·00890939	705	5·41	
		670		6·20
11	0·00988325	636	5·48	
		608		5·99
12	0·01082315	581	5·52	
		558		5·72
13	0·01172077	536	5·53	
15	0·01337545	469	5·50	
		457		4·87
16	0·01413967	444	5·46	
		433		4·67
17	0·01487227	422	5·41	
20	0·01694783	371	5·27	
		364		4·24
21	0·01761293	357	5·22	
		350		4·18
22	0·01826859	344	5·17	
25	0·02017568	311	5·04	
		307		4·04
26	0·02081038	302	5·00	
30	0·02329257	270	4·87	
		266		3·89
31	0·02390364	263	4·83	
40	0·02932744	214	4·61	
		212		3·82
41	0·02992771	210	4·59	
60	0·04136692	152	4·36	
		151		3·86
61	0·04197302	150	4·35	
100	0·06580556	95	4·17	
		95		3·90
101	0·06641823	95	4·17	
200	0·1258999	50	4·00	
		50		3·71
201	0·1264826	50	4·00	

Table 8

Earth flattening approximation for Jeffreys–Bullen Model.
σ—frequency, T—period, C—phase velocity, U—group velocity
$n(n+1) = a^2k^2$

n	σ	T	C	U
	s^{-1}	s	km/s	km/s
15	0·0150400	418	6·18	
		408		4·62
16	0·0157654	399	6·09	
20	0·0184944	340	5·75	
		334		4·15
21	0·0191461	328	5·67	
30	0·0246208	255	5·14	
		252		3·76
31	0·0252110	249	5·10	
40	0·0304432	206	4·79	
		204		3·69
41	0·0310228	203	4·76	
60	0·0421465	149	4·44	
		148		3·77
61	0·0427380	147	4·43	
100	0·0660051	95	4·18	
		94		3·81
102	0·0672006	93	4·18	
200	0·125091	50	3·97	
		50		3·70
201	0·125672	50	3·97	
400	0·234692	27	3·73	
		27		3·25
402	0·235712	27	3·73	
600	0·331074	19	3·51	
		19		2·96
604	0·332931	19	3·51	

Table 9

Period T, phase velocity C and group velocity U for Gutenberg Model obtained by the normal mode solution. n denotes the order of spherical harmonic. The interpolation is based on the exact values, computed at intervals of 5. For $n < 17$, the solution extends through the mantle and core; for $n > 17$ the amplitude in the core is negligible.

	T min		C	U		T min		C	U
n	Exact	Interpol.	km/s	km/s	n	Exact	Interpol.	km/s	km/s
2	53·52		4·99		33		4·06	4·91	
3	35·33		5·40	6·82	34		3·97	4·87	
4	25·54		5·80	7·52	35	3·89		4·84	
5	19·66		6·17	7·89	36		3·81	4·80	
6	15·92		6·45	7·85	37		3·73	4·77	
7	13·44		6·62	7·47	38		3·66	4·74	
8	11·74		6·69	6·84	39		3·59	4·71	
9	10·54		6·66	6·16	40	3·52		4·68	
10	9·65		6·59	5·61			3·49		3·58
11	8·95	8·93	6·48	5·23	41	3·45	3·45	4·65	
12	8·38	8·35	6·37	4·97	42		3·39	4·63	
13	7·90	7·87	6·26	4·80	43		3·33	4·61	
14	7·48	7·46	6·15	4·67	44		3·27	4·58	
15	7·11		6·05	4·56	45	3·22		4·56	
16	6·78	6·78	5·96	4·46	46		3·16	4·54	
17	6·49	6·49	5·87	4·35	47		3·11	4·52	
18	6·23	6·23	5·79	4·26	48		3·06	4·50	
19	6·00	6·00	5·71	4·17	49		3·01	4·48	
20	5·78		5·63	4·09	50	2·96		4·46	3·61
21	5·59	5·59	5·56		51		2·91	4·45	
22		5·40	5·49		52		2·87	4·43	
23		5·24	5·42		53		2·82	4·42	
24	5·08	5·08	5·36		54		2·78	4·40	
		5·01		3·82	55	2·74		4·39	3·66
25	4·94		5·30		56		2·70	4·38	
26		4·81	5·24		57		2·66	4·36	
27		4·68	5·18		58		2·62	4·35	
28		4·56	5·13		59		2·58	4·34	
29		4·45	5·08		60	2·55		4·33	
30	4·34		5·04				2·53		3·67
	4·29			3·64	61	2·51		4·32	
31	4·24	4·24	4·99						
32		4·15	4·95						

Table 10

Earth flattening approximation for Gutenberg Model
σ—frequency, T—period, C—phase velocity, U—group velocity
$n(n+1) = a^2k^2$

n	σ	T	C	U
	s^{-1}	s	km/s	km/s
10	0·0112644	558	6·83	
		535		6·52
11	0·0122873	511	6·81	
		494		5·81
12	0·0131994	476	6·73	
		462		5·19
13	0·0140140	448	6·61	
		437		4·73
14	0·0147563	426	6·48	
		416		4·42
15	0·0154503	407	6·35	
		398		4·22
16	0·0161131	390	6·22	
20	0·0186168	338	5·79	
		332		3·86
21	0·0192227	327	5·70	
30	0·0244724	257	5·11	
		254		3·62
31	0·0250403	251	5·06	
40	0·0300884	209	4·73	
		207		3·57
41	0·0306482	205	4·71	
60	0·0414016	152	4·36	
		151		3·66
61	0·0419763	150	4·35	
100	0·0649582	97	4·12	
		96		3·34
101	0·0655616	96	4·12	
200	0·126078	50	4·01	
		50		3·89
201	0·126689	50	4·01	

Table 11

Plane Earth approximation for Gutenberg Model
σ—frequency, T—period, C—phase velocity, U—group velocity
$n(n+1) = a^2k^2$

n	σ	T	C	U
	s^{-1}	s	km/s	km/s
12	0·0112840	557	5·75	
		539		6·30
13	0·0122726	512	5·79	
		494		5·97
14	0·0132103	476	5·80	
		461		5·57
15	0·0140852	446	5·79	
		434		4·75
16	0·0148923	422	5·75	
		411		3·95
17	0·0156372	401	5·69	
20	0·0176298	356	5·48	
		350		3·95
21	0·0182502	344	5·41	
25	0·0206441	304	5·16	
		300		3·73
26	0·0212293	296	5·10	

of gravitational forces in the Earth flattening approximation makes but a small improvement, as shown by the points ⊙.

A comparison of the theoretical group velocity curves of Rayleigh waves for various models with the observations of Ewing & Press (1956) is shown in Figures 9, 10 and 11. The theoretical group velocity curves for Bullen's Model B, as well as for the "Jeffreys–Bullen" Model, shown in Figures 9 and 10, are distinctly higher than the observed values for periods greater than 50s. On the other hand, the theoretical group velocity curve for the Gutenberg Model agrees with the observations for periods up to about 380s. This agreement can be taken as evidence in favour of Gutenberg's low-velocity layer as was first demonstrated by Takeuchi, Press & Kabayashi (1959), and by Dorman, Ewing & Oliver (1960).

The few observed Rayleigh wave group velocities in the high-period region of 400 to about 630s are much lower than the theoretical values for all the three models. It would be interesting to substantiate this by further observations. The observed phase velocities C of Nafe and Brune shown in Figure 11 agree very well with the theoretical normal mode solution for $T < 300$. For $T > 300$s, the observed C values are higher than the theoretical.

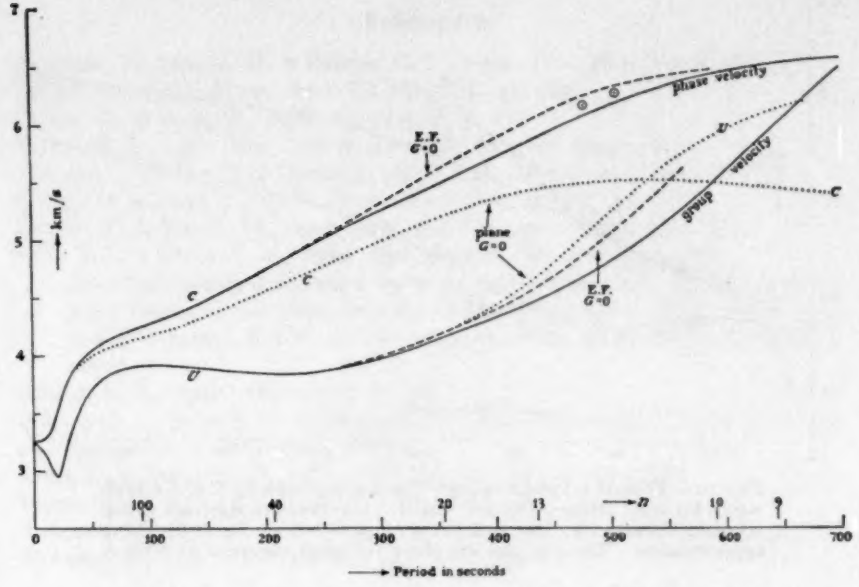


FIG. 8.—Group velocity U and phase velocity C for Bullen's model B. ———— exact normal mode theory, – – – – Earth flattening approximation with neglect of gravity, ... plane Earth with neglect of gravity, ⊙ Earth flattening approximation with gravity included.

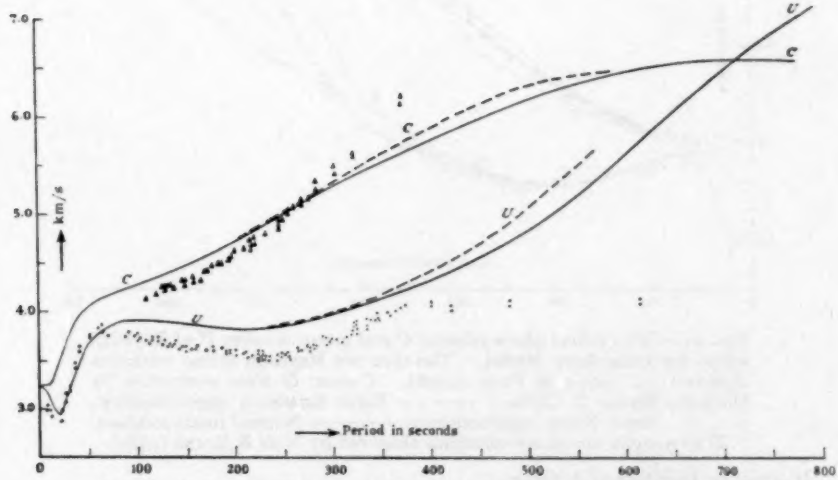


FIG. 9.—Theoretical group velocity U and phase velocity C of Rayleigh waves for Bullen's model. The dots are Rayleigh wave group velocities observed by Ewing & Press (1956). ———— normal mode solution, – – – – – Earth flattening approximation. The triangles are phase velocities observed by Nafe & Brune (1960).

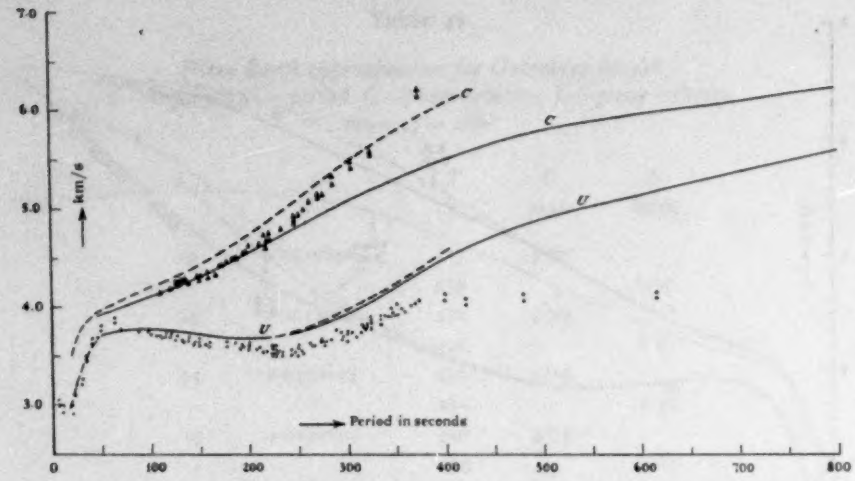


FIG. 10.—Theoretical phase velocity C and group velocity U of Rayleigh waves for the "Jeffreys–Bullen" model. The dots are Rayleigh group velocities observed by Ewing & Press (1956) – – – – – Earth flattening approximation. The triangles are phase velocities observed by Nafe & Brune (1960).

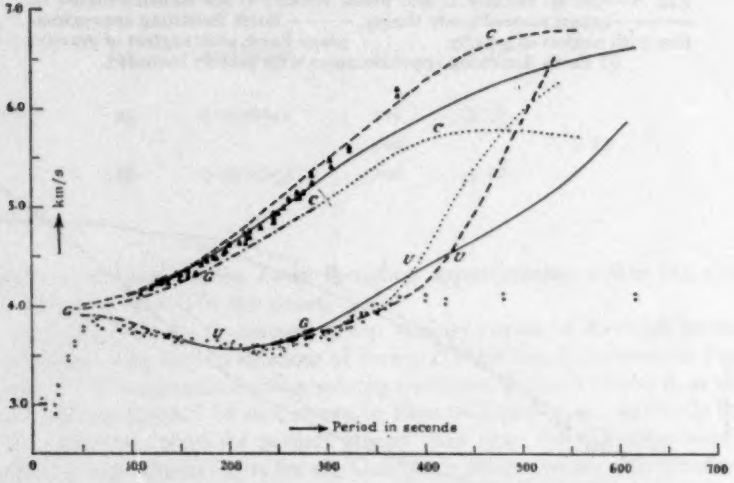


FIG. 11.—Theoretical phase velocity C and group velocity U of Rayleigh waves for Gutenberg Model. The dots are Rayleigh group velocities observed by Ewing & Press (1956). Curves G were computed by Dorman, Ewing & Oliver. – – – – – Earth flattening approximation, ... Plane Earth approximation, ——— Normal mode solution. The triangles are phase velocities observed by Nafe & Brune (1960).

Weizmann Institute of Science,
Rehovot,
Israel:

1960 *September*

References

Alterman, Z., Jarosch, H. & Pekeris, C. L., 1959. *Proc. Roy. Soc. A*, **252**, 80.
Benioff, H., 1954. *Trans. Amer. Geophys. Un.*, **35**, 985.
Benioff, H. & Press, F., 1958. *Geophys. J.*, **1**, 208.
Bullen, K. E., 1950. *Mon. Not. R. Astr. Soc. Geophys. Suppl.*, **6**, 50.
Dorman, J., Ewing, M. & Oliver, J., 1960. *Bull. Seismol. Soc. Amer.*, **50**, 87.
Ewing, M. & Press, F., 1956. *Trans. Amer. Geophys. Un.*, **37**, 213.
Koo, Y. C. & Katzin, M., 1960. *Jour. Res. Nat. Bur. Stand.*, **64D.**, 61.
Nafe, J. E. & Brune, J. N., 1960. *Bull. Seismol. Soc. Amer.*, **50**, July.
(In a forthcoming publication the authors re-analyse their data and arrive at lower values for the phase velocity. The reduction amounts to about 1 per cent at a period of 200 s and increases to about 3 per cent at a period of 400 s.)
Pekeris, C. L., 1946. *Phys. Rev.*, **70**, 518.
Pekeris, C. L., Jarosch, H. & Alterman, Z., 1959. *Proceedings, Third International Symposium on Earth Tides, Trieste.*
Press, F. & Takeuchi, H., 1960. *Bull. Seismol. Soc. Amer.*, **50**, 81.
Pryce, M. H. L., 1953. *Advances in Physics*, **2**, 67.
Takeuchi, H., Press, F. & Kobayashi, N., 1959. *Bull. Seismol. Soc. Amer.*, **49**, 355.
Stoneley, R., 1953. *Bull. Seismol. Soc. Amer.*, **43**, 127.

Calcul Approché de la Periode des Oscillations Spheroïdales de la Terre

Nelly Jobert

"Step after step the ladder is ascended."—*S. Herbert*, Jacula Prudentum.

Résumé

Dans la première partie on étudie les oscillations sphéroïdales en l'absence de gravitation. On considère l'effet du noyau terrestre, et on obtient la solution exacte pour une sphère à noyau homogène liquide et manteau homogène solide, ayant respectivement les propriétés moyennes du noyau et du manteau terrestres. Par application du principe de Rayleigh suivant une méthode dûe à H. Jeffreys, on trouve une valeur approchée de la période de vibration pour le modèle (A) de K. Bullen.

Dans la deuxième partie on tient compte des forces de gravitation, dont on étudie l'effet sur des modèles de complexité croissante. Par application du principe de Rayleigh on obtient une valeur approchée de la période des oscillations sphéroïdales pour les modèles (A) et (B) de K. Bullen.

Introduction

Les plus grandes périodes de vibration propre d'une sphère élastique homogène sont celles où la fonction sphérique représentant la répartition du déplacement en surface, a pour indice $n = 2$. Le calcul de ces périodes a été entrepris depuis longtemps; une bibliographie très complète de la question a été publiée par F. Gilbert & G. J. F. MacDonald (1960). H. Lamb (1882) considère deux classes de vibrations: les oscillations de 1ère classe, ou oscillations de torsion, dues aux ondes SH (pour $n = 2$, le mouvement de la sphère est alors une torsion de signe opposé dans les deux hémisphères), et les oscillations de 2ème classe, dues à la fois aux ondes SV et aux ondes P (pour $n = 2$, la sphère prend alternativement la forme d'un ellipsoïde allongé puis aplati, d'où le nom d'oscillations sphéroïdales). Les forces de gravitation n'interviennent pas dans les oscillations de 1ère classe. Leur effet sur celles de 2ème classe a été étudié par différents auteurs; en particulier, A. E. H. Love (1911) a étudié le cas d'une sphère homogène, compressible (coefficient de Poisson $\sigma = \frac{1}{4}$) ayant la rigidité de l'acier et une masse égale à celle de la Terre.

Ces travaux, purement théoriques, ont pris un intérêt nouveau avec l'observation, lors du séisme du Kamchatka du 4 novembre 1952, d'une onde très longue dont la période est voisine de 57 minutes (H. Benioff 1954) et qui est donc attribuable à une oscillation sphéroïdale plutôt qu'à une oscillation de torsion, dont la période est plus courte. La comparaison d'une telle observation avec les résultats

théoriques est d'un grand intérêt; malheureusement l'observation directe des vibrations propres de la Terre est extrêmement rare. On peut espérer que les progrès accomplis en matière d'analyse permettront de les mettre en évidence.

Le calcul de la période des oscillations sphéroïdales, pour la Terre, a été entrepris tout d'abord d'une façon approchée par application du principe de Rayleigh (en effet, dès qu'on s'écarte d'un modèle homogène, les équations différentielles du mouvement ont une complexité telle qu'on ne peut aborder le calcul exact qu'à l'aide d'un calculateur électronique). Le principe de Rayleigh donne une valeur trop faible de la période, avec une erreur d'autant plus petite que la solution d'essai choisie est plus voisine de la vraie solution. R. Stoneley (1926) a proposé une méthode variationelle basée sur le principe de Ritz–Rayleigh, qui a été utilisée par C. L. Pekeris & H. Jarosch (1958). Nous avons utilisé une autre méthode d'application du principe de Rayleigh, analogue à celle que H. Jeffreys avait appliquée en séismologie (H. Jeffreys 1935, R. Stoneley 1953): le calcul approché de la dispersion des ondes de surface pour un modèle complexe peut être fait de manière simple en prenant pour solution d'essai la solution exacte pour un modèle plus simple (ainsi pour les ondes de Rayleigh à la surface d'un demi-espace recouvert d'une ou deux couches superficielles, la solution d'essai est celle pour un demi-espace homogène). Les calculs exposés ici ont conduit à des résultats publiés antérieurement dans des Notes très succinctes (1957). Depuis un résultat exact a été obtenu (Z. Alterman, H. Jarosch & C. L. Pekeris 1959) qui nous permettra une comparaison.

I. Oscillations de 2$^{\text{ème}}$ classe en l'absence de gravitation

Considérons un point M à l'intérieur d'une sphère élastique de rayon a. Les coordonnées sphériques de M sont: r, θ, ϕ. Soit ρ la densité au point M, λ et μ les coefficients de Lamé définissant les propriétés élastiques en M. Soit $\mathscr{W} = (\mu/\rho)^{\frac{1}{2}}$ la vitesse des ondes S, $\mathscr{V} = [(\lambda+2\mu)/\rho]^{\frac{1}{2}}$ celle des ondes P. Considérons un mouvement sinusoïdal en fonction du temps, de période $T = 2\pi/p$. Soit $\mathbf{d} \cdot \cos pt$ le déplacement en M, que nous supposerons indépendant de ϕ (symétrie de révolution).

1. Sphère homogène

Les équations du mouvement dans un milieu homogène s'écrivent:

$$(\lambda+\mu)\ \text{grad}\,\Delta + \mu\nabla^2\mathbf{d} + \rho p^2\mathbf{d} = 0 \qquad (1)$$

$\Delta = \text{div}\,\mathbf{d}$ représente la dilatation,
$\nabla^2\mathbf{d}$ est le vecteur ayant pour composantes cartésiennes les Laplaciens de chaque composante de $\mathbf{d}$. Soit u la composante méridienne de $\mathbf{d}$, w la composante radiale, Γ la composante suivant ϕ du vecteur curl $\mathbf{d}$:

$$\left.\begin{aligned} \Gamma = (\text{curl}\,\mathbf{d})_\phi &= \frac{1}{r}\frac{\partial}{\partial r}(ru) - \frac{1}{r}\frac{\partial w}{\partial \theta} \\ \Delta = \text{div}\,\mathbf{d} &= \frac{1}{r^2}\frac{\partial}{\partial r}(wr^2) + \frac{1}{r\sin\theta}\frac{\partial}{\partial \theta}(u\sin\theta) \end{aligned}\right\}. \qquad (2)$$

En prenant successivement la divergence et le rotationnel (curl) du premier membre de (1), il vient:

$$(\lambda+2\mu)\nabla^2\Delta+\rho p^2\Delta = 0$$
$$\mu\nabla^2\Gamma+\rho p^2\Gamma = 0.$$

Nous chercherons des solutions de la forme:

$$\Delta = \delta(r)P_n(\cos\theta),$$
$$\Gamma = \gamma(r)P_n^1(\cos\theta)$$

où P_n est la fonction sphérique d'ordre n, P_n^1 sa dérivée par rapport à θ. Soit: $h^2 = \rho p^2/(\lambda+2\mu)$, $k^2 = \rho p^2/\mu$; $\delta(r)$ et $\gamma(r)$ doivent satisfaire respectivement les équations:

$$d^2\delta/dr^2+(2/r)\,d\delta/dr+[h^2-n(n+1)r^{-2}]\delta = 0$$
$$d^2\gamma/dr^2+(2/r)\,d\gamma/dr+[k^2-n(n+1)r^{-2}]\gamma = 0.$$

La solution finie au centre est de la forme:

$$\delta = M(hr)^{-\frac{1}{2}}J_{n+\frac{1}{2}}(hr)$$
$$\gamma = N(kr)^{-\frac{1}{2}}J_{n+\frac{1}{2}}(kr),$$

$J_{n+\frac{1}{2}}$ étant la fonction de Bessel de première espèce, d'ordre $n+\frac{1}{2}$. Prenons les composantes du déplacement sous la forme:

$$u = U(r)P_n^1(\cos\theta), \qquad w = W(r)P_n(\cos\theta).$$

D'aprés les équations (2) la solution finie au centre pour le déplacement est de la forme:

$$\left.\begin{aligned} rU &= A_1(hr)^{-\frac{1}{2}}J_{n+\frac{1}{2}}(hr)+B_1\partial[r(kr)^{-\frac{1}{2}}J_{n+\frac{1}{2}}(kr)]/\partial r \\ rW &= A_1r\partial[(hr)^{-\frac{1}{2}}J_{n+\frac{1}{2}}(hr)]/\partial r+n(n+1)B_1(kr)^{-\frac{1}{2}}J_{n+\frac{1}{2}}(kr) \end{aligned}\right\}$$

soit en posant:

$$\left.\begin{aligned} h_n(r) &= A_1(hr)^{-\frac{1}{2}}J_{n+\frac{1}{2}}(hr), \qquad h^0{}_{n-1}(r) = A_1(hr)^{\frac{1}{2}}J_{n-\frac{1}{2}}(hr) \\ k_n(r) &= -B_1(kr)^{-\frac{1}{2}}J_{n+\frac{1}{2}}(kr), \quad k^0{}_{n-1}(r) = -B_1(kr)^{\frac{1}{2}}J_{n-\frac{1}{2}}(kr) \\ rU &= h_n-(k^0{}_{n-1}-nk_n), \\ rW &= h^0{}_{n-1}-(n+1)h_n-n(n+1)k_n \end{aligned}\right\} \tag{3}$$

et

$$\delta = -h^2h_n, \qquad \gamma = k^2k_n.$$

Les conditions aux limites s'écrivent, pour $r = a$:

$$\left.\begin{aligned} \lambda\delta+2\mu\,dW/dr &= 0 \\ r^2d(U/r)/dr+W &= 0 \end{aligned}\right\} \tag{4}$$

soit dans le cas où $\lambda = \mu$ ($\sigma = \frac{1}{4}$):

$$\left.\begin{aligned} 3a\delta+2n(n+1)U-4W &= 0 \\ a\gamma+2W-2U &= 0. \end{aligned}\right\}$$

L'élimination des constantes A_1 et B_1 entre ces deux équations donne l'équation de dispersion:

$$[2n(n+1)-k^2a^2-4-4X][2n(n+1)-k^2a^2-4-2Y] = 4n(n+1)XY \qquad (5)$$

où:

$$X = a\frac{\partial}{\partial a}\log[(ha)^{-\frac{1}{2}}J_{n+\frac{1}{2}}(ha)]$$

$$Y = a\frac{\partial}{\partial a}\log[(ka)^{-\frac{1}{2}}J_{n+\frac{1}{2}}(ka)].$$

Λ étant la longueur d'onde et C la vitesse de phase, posons:

$$f = (n+\tfrac{1}{2})/a = 2\pi/\Lambda$$
$$r = (f^2-h^2)^{\frac{1}{2}} = f(1-C^2V^{-2})^{\frac{1}{2}}$$
$$s = (f^2-k^2)^{\frac{1}{2}} = f(1-C^2W^{-2})^{\frac{1}{2}}.$$

Lorsque n est grand, les expressions asymptotiques de Debye (Watson 1944, p. 243), permettent d'écrire:

$$X = a[r+O(a^{-1})]$$
$$Y = a[s+O(a^{-1})].$$

D'autre part:

$$n(n+1) = (n+\tfrac{1}{2})^2-\tfrac{1}{4}.$$

En négligeant les termes d'ordre a^{-2} dans l'équation (5), il vient:

$$(2f^2-k^2-4r/a)(2f^2-k^2-2s/a) = 4f^2rs.$$

Lorsque la longueur d'onde Λ diminue, les terms r/a et s/a deviennent négligeables et on retrouve l'équation donnant la vitesse de l'onde de Rayleigh à la surface d'un demi-espace homogène (K. Sezawa 1929).

Dans le cas $n = 2$, nous avons fait le calcul à l'aide des Tables du National Bureau of Standards; la plus petite racine de (5) est: $ka = 2{\cdot}639$ (H. Lamb 1882; A. E. H. Love 1944). On trouve:

$$B_1/A_1 = 0{\cdot}4807.$$

Les composantes de l'amplitude du déplacement correspondant sont représentées sur la Figure 2. Leur utilisation comme solution d'essai dans l'application du principe de Rayleigh n'a pas donné un bon résultat: la discontinuité due au noyau terrestre a trop d'importance pour les oscillations sphéroïdales.

2. Effet du noyau terrestre

Nous avons considéré tout d'abord un noyau vide ayant la dimension du noyau terrestre (rayon $b = 3\,473$ km, le rayon de la surface étant $a = 6\,371$ km). Dans le manteau, l'amplitude du déplacement cherché a pour composantes:

$$rU = g_2-l_1^0+2l_2,$$
$$rW = g_1^0-3g_2-6l_2, \qquad (6)$$

avec:

$$\delta = -h^2 g_2 \qquad \gamma = k^2 l_2$$
$$g_2(r) = A(hr)^{-\frac{1}{2}} J_{\frac{1}{2}}(hr) + \bar{A}(hr)^{-\frac{1}{2}} J_{-\frac{1}{2}}(hr)$$
$$l_2(r) = B(kr)^{-\frac{1}{2}} J_{\frac{1}{2}}(kr) + \bar{B}(kr)^{-\frac{1}{2}} J_{-\frac{1}{2}}(kr)$$
$$g_1^0(r) = A(hr)^{\frac{1}{2}} J_{\frac{1}{2}}(hr) - \bar{A}(hr)^{\frac{1}{2}} J_{-\frac{1}{2}}(hr)$$
$$l_1^0(r) = B(kr)^{\frac{1}{2}} J_{\frac{1}{2}}(kr) - \bar{B}(kr)^{\frac{1}{2}} J_{-\frac{1}{2}}(kr).$$

Les conditions aux limites sont les mêmes sur le noyau qu'à la surface, ce sont les conditions (4). L'élimination des quatre constantes entre les quatre équations donne l'équation aux périodes propres, qui se trouve satisfaite pour:

$$ka = 1{\cdot}81.$$

Les valeurs de ka correspondant à diverses valeurs du rayon du noyau vide ont été portées sur la Figure 1. Pour une coque infiniment mince, on a: $ka = 1{\cdot}176$

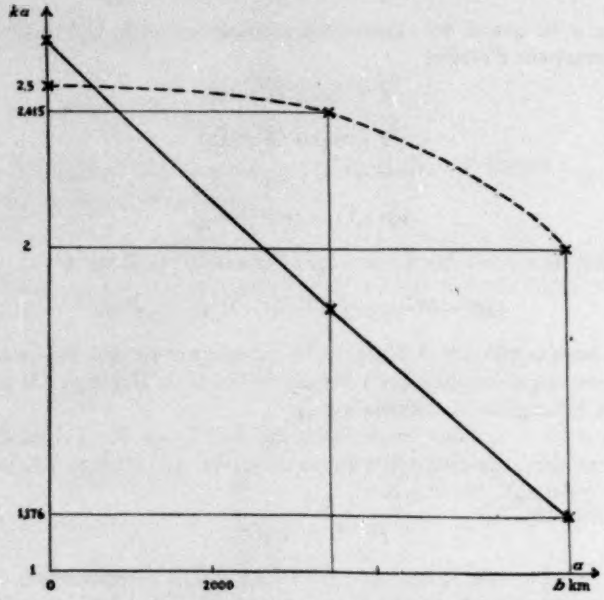


FIG. 1.—Variation de ka avec le rayon b du noyau.
– – – – – oscillations de 1ère classe.
——— oscillations de 2ème classe.

(Rayleigh 1894). Sur la même figure ont été portées les valeurs de ka correspondantes pour les vibrations de 1ère classe de même indice, $n = 2$. On voit que l'effet de la discontinuité du noyau est beaucoup plus faible pour ces dernières. L'amplitude des oscillations de 1ère classe décroît en effet plus vite avec la profondeur que celle des oscillations de 2ème classe: pour une sphère homogène le déplacement au voisinage du centre est d'ordre r^{n-1} pour la 2ème classe, r^n pour la 1ère. Il en résulte que, au moins pour les petites valeurs de n, les ondes de Rayleigh sont

influencées par des structures plus profondes que les ondes de Love de même longueur d'onde.

Nous avons ensuite considéré une sphère à noyau liquide homogène et à manteau élastique homogène, ayant respectivement les propriétés moyennes du noyau et du manteau terrestres. Ce modèle est celui utilisé par T. Matumoto et Y. Satô (1954) pour les oscillations de 1ère classe:

$$\lambda_c/\lambda_m = 4{\cdot}918, \qquad \rho_c/\rho_m = 2{\cdot}4, \qquad \mathscr{V}_c = 9{\cdot}3\,\text{km/s}, \qquad \mathscr{W}_c = 0$$
$$\mathscr{W}_m = 6{\cdot}5\,\text{km/s}, \qquad \lambda_m = \mu_m.$$

Dans le manteau le déplacement a la forme donnée par (6), dans le noyau celle donnée par (3) où ne figurent pas de fonctions de kr puisque $\mu_c = 0$. Les conditions aux limites sont: à la surface ($r = a$) les deux conditions (4) auxquelles s'ajoutent pour la surface du noyau ($r = b$):

$$\lambda_c\delta_c = \lambda_m\delta_m + 2\mu_m dW_m/dr$$
$$r^2\,d(U_m/r)/dr + W_m = 0$$
$$W_m = W_c.$$

Ces 5 équations sont compatibles pour:

$$ka = 1{\cdot}445.$$

La période de vibration est alors:

$$T = 71\ \text{min}.$$

Les constantes des expressions du déplacement ont pour valeur:

$$A_m = -232{\cdot}2 \qquad B_m = -32{\cdot}61 \qquad A_c = 107{\cdot}7$$
$$\bar{A}_m = 0{\cdot}3982 \qquad \bar{B}_m = -1.$$

Les composantes du déplacement ont été représentées sur la Figure 2. Elles seront utilisées comme solution d'essai (solution 1) dans les applications du principe de Rayleigh.

3. Application du principe de Rayleigh au modèle (A) de K. Bullen (1953)

On obtient une valeur approchée par défaut de la période de vibration en égalant l'énergie potentielle élastique moyenne, pour une période et pour toute la sphère, à l'énergie cinétique moyenne. Les expressions utilisées sont les suivantes:

Energie cinétique moyenne:

$$\mathscr{T}_0 = (2T)^{-1}\int_0^T dt\,(2\pi)^{-1}\int_0^{2\pi} d\phi \int_0^{\pi} \sin\theta\,d\theta \int_0^a [U^2(P_n^1)^2 + W^2(P_n)^2]p^2r^2\sin^2pt\,dr$$

soit

$$\mathscr{T}_0 = p^2 I_0/(4n+2). \tag{7}$$

Pour $n = 2$:

$$I_0 = \int_0^a (r^2W^2 + 6r^2U^2)\,dr.$$

Energie potentielle moyenne:
De la même façon on trouve:

$$\mathscr{E}_0 = (4n+2)^{-1}\int_0^a [(\lambda\delta+2\mu W')W'+\lambda\delta r^{-1}\{2W-n(n+1)U\}+ \\ +\mu r^{-2}\{2(2W^2+n(n+1)(n^2+n-1)U^2-2n(n+1)UW)+ \\ +n(n+1)(rU'-U+W)^2\}]r^2\,dr. \tag{8}$$

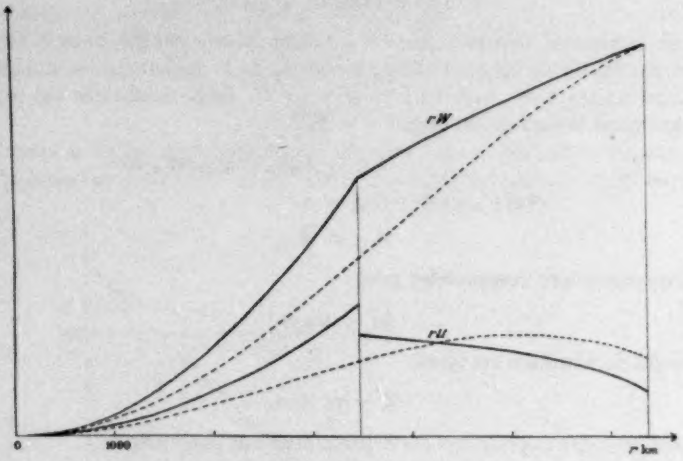


FIG. 2.—Variation des composantes du déplacement avec le rayon pour les oscillations de 2ème classe en l'absence de gravitation ($n = 2$).
——— Solution I: sphère à manteau et noyau liquide homogènes (Modèle de Matumoto et Satô).
– – – – – Sphère homogène.

Pour $n = 2$, le calcul est fait à l'aide de l'expression:

$$\mathscr{E}_0 = (10)^{-1}\int_0^a [\lambda r^2\delta^2+2\mu(r\delta-2W+6U)^2+4\mu\{6U^2+(W-3U)^2\}+ \\ +6\mu(r\gamma+2W-2U)^2]\,dr. \tag{8'}$$

Application numérique:

La solution d'essai est la solution 1. Les intégrales ont été évaluées graphiquement. L'énergie potentielle du noyau est extrêmement faible, par contre son énergie cinétique représente environ le tiers de l'énergie cinétique totale. On trouve:

$$T = 66{\cdot}3 \text{ min.}$$

II. Influence des forces de gravitation

1. Sphère homogène

Le problème a été traité par Love (1911) pour une sphère homogène compressible. Nous établirons d'abord la forme du déplacement; le même calcul servira

plus loin pour la construction d'une solution d'essai. On obtient les équations du mouvement en ajoutant à la force élastique, une force due aux effets de gravitation mutuelle entre les points de la sphère en mouvement.

Soit $V_0 = (\frac{3}{8})\pi G\rho(3a^2-r^2)$, le potentiel de gravitation en l'absence d'oscillations, $G = 6{,}67\ 10^{-8}$ cgs étant la constante de gravitation. Soit $v\cos pt$, avec $v = V(r)P_n(\cos\theta)$, le potentiel supplémentaire dû aux changements dans la répartition des masses à l'intérieur de la sphère lors d'une oscillation. Au repos l'état initial des tensions compense exactement la force de gravité. La force agissant sur un élément de volume en mouvement se compose de trois termes; l'un vient de ce que la force de gravité au point déplacé ne compense plus exactement la tension initiale, le deuxième provient du changement local de densité, et le troisième est un effet du potentiel supplémentaire. Les équations du mouvement s'écrivent alors:

$$(\lambda+\mu)\operatorname{grad}\Delta+\mu\nabla^2\mathbf{d}+\rho\operatorname{grad}(W\,dV_0/dr)-\rho\Delta\operatorname{grad}V_0+\rho\operatorname{grad}v = -\rho p^2\mathbf{d}, \tag{9}$$

ou encore:

$$\left.\begin{aligned}(\lambda+2\mu)\,d\delta/dr+n(n+1)\mu\gamma/r-4\pi G\rho^2\,d(rW)/3dr+4\pi G\rho^2\delta r/3+\rho\,dV/dr\\ = -\rho p^2W\\ (\lambda+2\mu)\delta+\mu\,d(r\gamma)/dr-4\pi G\rho^2Wr/3+\rho V = -\rho p^2rU.\end{aligned}\right\} \tag{9'}$$

En procédant comme avec l'équation (1) et en utilisant la relation:

$$\nabla^2v = 4\pi G\rho\Delta, \tag{10}$$

il vient:

$$\left.\begin{aligned}(\lambda+2\mu)[r^2d^2(\delta r)/dr^2-n(n+1)r\delta]+r^3\delta(\rho p^2+16\pi G\rho^2/3)\\ = 4\pi Gn(n+1)r^3\gamma\rho^2/3\\ \mu[r^2d^2(r\gamma)-n(n+1)r\gamma]+\rho p^2r^3\gamma = 4\pi Gr^3\delta\rho^2/3.\end{aligned}\right\}$$

Soit $g = 4\pi G\rho a/3$ l'intensité de la pesanteur à la surface. $r\gamma$ et $r\delta$ satisfont l'équation différentielle du 4ème ordre:

$$[d^2/dr^2-n(n+1)r^{-2}+\alpha^2][d^2/dr^2-n(n+1)r^{-2}+\beta^2]r\gamma = 0 \tag{11}$$

où $\pm\alpha$ et $\pm\beta$ sont les racines de l'équation du 4ème degré:

$$\begin{aligned}\mu(\lambda+2\mu)x^4-[\mu(\rho p^2+4g\rho/a)+(\lambda+2\mu)\rho p^2]x^2+\rho p^2(\rho p^2+4g\rho/a)-\\ -n(n+1)\rho^2g^2a^{-2} = 0.\end{aligned} \tag{12}$$

La solution générale, finie au centre, de l'équation (11) est de la forme:

$$r^{\frac{1}{2}}[MJ_{n+\frac{1}{2}}(\alpha r)+NJ_{n+\frac{1}{2}}(\beta r)].$$

Si l'on écrit:

$$-3g\delta/a = A\alpha^2(\alpha r)^{-\frac{1}{2}}J_{n+\frac{1}{2}}(\alpha r)+B\beta^2(\beta r)^{-\frac{1}{2}}J_{n+\frac{1}{2}}(\beta r)$$

il vient d'après (10):

$$V = A(\alpha r)^{-\frac{1}{2}}J_{n+\frac{1}{2}}(\alpha r)+B(\beta r)^{-\frac{1}{2}}J_{n+\frac{1}{2}}(\beta r)+C(r/a)^n.$$

Posons:

$$\begin{aligned}(1-n\eta)h_n(r) &= (Aa/3g)(\alpha r)^{-\frac{1}{2}}J_{n+\frac{1}{2}}(\alpha r)\\ n\epsilon k_n(r) &= (Ba/3g)(\beta r)^{-\frac{1}{2}}J_{n+\frac{1}{2}}(\beta r)\end{aligned}$$ (et de même pour h^0, k^0).

Les composantes de l'amplitude du déplacement ont alors la forme:

$$\left.\begin{aligned} rU &= h_n-(k^0_{n-1}-nk_n)+H(r)\\ rW &= h^0_{n-1}-(n+1)h_n-n(n+1)k_n+nH_n(r)\end{aligned}\right\} \qquad (13)$$

avec:

$$H(r) = \epsilon k^0_{n-1}(r)-\eta h^0_{n-1}(r)+C'(r/a)^n, \qquad C' = C/(-p^2+ng/a)$$
$$\epsilon = (\mu\beta^2-\rho p^2)/[\mu\beta^2-\rho p^2+n\rho g/a]$$
$$\eta = (\rho g/a)/[\mu\alpha^2-\rho p^2+n\rho g/a].$$

Ces expressions peuvent être identifiées avec celles de Love (1911) (p. 130) et de Pekeris (1958, p. 186) pour $n = 2$. Pour de grandes valeurs de n lorsque la période diminue, $p^2 \gg g/a$, les racines α et β de l'équation (12) tendent respectivement vers h et k, et ϵ, η et C' tendent vers zéro. Le déplacement tend vers la forme donnée par les équations (3).

Les conditions à la surface sont les conditions (4) auxquelles s'ajoute, pour le potentiel, la condition:

$$dV/dr-V(a)a^3\,d(r^{-3})/dr = 3gW/a, \qquad \text{pour } r = a.$$

Pour une sphère homogène ayant la rigidité de l'acier, la masse et les dimensions de la Terre, Love (1911) a trouvé une période de 60 min pour l'oscillation sphéroïdale, soit une diminution de 6 min par suite de l'effet de gravité. Le même calcul a été fait par Pekeris (1958), pour un modèle ayant les propriétés moyennes de la Terre. La période trouvée est de 44·3 min.

2. Sphère à noyau liquide et à manteau solide homogène (Modèle de Matumoto et Satô (1954))

(*a*) *Application du principe de Rayleigh, la solution d'essai étant la solution pour le même modèle en l'absence de gravitation* (*Solution I*).

L'énergie cinétique moyenne sur toute la sphère pendant une période, est donnée par l'expression (7). L'énergie potentielle élastique moyenne est donnée par l'expression (8). En transformant cette dernière, suivant la méthode utilisée par Jeffreys (1935) pour un demi-espace homogène, il vient:

$$(4n+2)\,\mathscr{E}_0 = [(\lambda\delta+2\mu W')Wr^2+\mu n(n+1)(rU'-U+W)rU]_0^a-$$
$$-\int_0^a [\{(\lambda+2\mu)\delta' r+\mu n(n+1)\gamma\}rW+$$
$$+n(n+1)\{(\lambda+2\mu)\delta+\mu(r\gamma)'\}rU]\,dr.$$

Le terme intégré est nul si les conditions aux limites sont satisfaites, et l'intégrale peut être transformée en utilisant les équations du mouvement. Comme la solution d'essai I satisfait les conditions aux limites et les équations du mouvement en l'absence de gravitation, on en déduit:

$$\mathscr{E}_0 = \mathscr{T}_0 = p_0^2 I_0/10$$

où

$$p_0 = 2\pi T_0,\ T_0 = 71 \text{ min}.$$

Si l'on tient compte des forces de gravitation, il faut ajouter à l'énergie potentielle élastique l'énergie de gravitation. Celle-ci $\mathscr{E}_g$ est calculée à partir de la solution d'essai. Le principe de Rayleigh donne la valeur approchée de p. Pour $n = 2$:

$$\mathscr{E}_g + \mathscr{E}_0 = p^2 I_0/10 \tag{14}$$

ou

$$(p/p_0)^2 - 1 = 10\mathscr{E}_g/(p_0^2 I_0). \tag{14'}$$

Calculons d'abord le potentiel de gravitation en l'absence d'oscillations, soit $V_0(r, \theta)$ en un point M de coordonnées $(r, \theta, 0)$ à l'intérieur de la sphère. Il peut être considéré comme la somme de deux potentiels:

V_1: potentiel d'une sphère homogène de densité ρ_m, de rayon a
V_2: potentiel d'une sphère homogène de densité $\rho_c - \rho_m$, de rayon b.

On a donc:

$$V_0 = V_1 + V_2 = \begin{cases} (2\pi G/3)[\rho_m(3a^2 - r^2) + 2(\rho_c - \rho_m)b^3/r] & \text{si } b < r < a \\ (2\pi G/3)[3a^2\rho_m + 3b^2(\rho_c - \rho_m) - \rho_c r^2] & \text{si } 0 < r < b. \end{cases}$$

Dans le calcul du potentiel supplémentaire, nous n'écrirons pas le facteur de temps $\cos pt$. Ce potentiel est dû à une distribution en volume $-\rho_c\Delta_c$ dans le noyau, $-\rho_m\Delta_m$ dans le manteau, et des distributions superficielles: $(\rho_c - \rho_m)w(b)$ à la surface du noyau, $\rho_m w(a)$ à la surface de la Terre.

Soit $d\tau$ un élément de volume entourant un point $M_1(r_1, \theta_1, \phi_1)$, $d\sigma$ un élément de surface. Soit D la distance MM_1:

$$D^2 = r^2 + r_1^2 - 2rr_1\cos\Theta$$

avec:

$$\cos\Theta = \cos\theta\cos\theta_1 + \sin\theta\sin\theta_1\cos\phi_1.$$

Au point M le potentiel supplémentaire est donné par:

$$v(r, \theta) = G[-\int_0^b\int_0^{2\pi}\int_0^{\pi} \rho_c\Delta_c D^{-1}\,d\tau + (\rho_c - \rho_m)\int_0^{2\pi}\int_0^{\pi} w(b)D^{-1}\,d\sigma_b$$

$$-\int_b^a\int_0^{2\pi}\int_0^{\pi} \rho_m\Delta_m D^{-1}\,d\tau + \rho_m\int_0^{2\pi}\int_0^{\pi} w(a)D^{-1}\,d\sigma_a.$$

En utilisant les relations:

$$D^{-1} = \sum r_1^m r^{-(m+1)}P_m(\cos\Theta) \qquad \text{si } r_1 < r$$
$$D^{-1} = \sum r^m r_1^{-(m+1)}P_m(\cos\Theta) \qquad \text{si } r < r_1$$

et en intégrant, on voit apparaître la fonction:

$$f(r, r_1) = r^2 r_1^{-3} \qquad \text{pour } r < r_1,$$
$$f(r, r_1) = r_1^2 r^{-3} \qquad \text{pour } r > r_1.$$

Nous écrirons le potentiel supplémentaire sous la forme:

$$v(r, \theta) = V(r)P_2(\cos\theta) = (4\pi G/5)\rho_m F(r)P_2(\cos\theta)$$

où:

$$F(r) = -(\rho_c/\rho_m)\int_0^b \delta_c f(r,r_1) r_1^2\, dr_1 + (-1+\rho_c/\rho_m) b^2 W(b) f(b,r) -$$

$$-\int_0^a \delta_m f(r,r_1) r_1^2\, dr_1 + a^2 W(a) f(a,r).$$

Les expressions de δ_c et δ_m sont données par (3) et (6); les intégrales qui les contiennent sont calculables à partir des relations:

$$\int^z z^{m+1} C_m(z)\, dz = z^{m+1} C_{m+1}(z), \qquad \int^z z^{-m+1} C_m(z)\, dz = -z^{-m+1} C_{m-1}(z).$$

(Watson 1944, p. 132), où C_m est une fonction de Bessel d'ordre m. On a donc:
—dans le noyau:

$$F_c(r) = (\rho_c/\rho_m)[5h_2(r) - (r/b)^2 h_1^0(b)] + (-1+\rho_c/\rho_m) b W(b)(r/b)^2 - g_1^0(a)(r/a)^2 + \\ + g_1^0(b)(r/b)^2 + aW(a)(r/a)^2$$

—dans le manteau:

$$F_m(r) = (\rho_c/\rho_m)(b/r)^3[5h_2(b) - h_1^0(b)] + (-1+\rho_c/\rho_m) b W(b)(b/r)^3 + 5g_2(r) - \\ - g_1^0(a)(r/a)^2 - (b/r)^3[5g_2(b) - g_1^0(b)] + aW(a)(r/a)^2.$$

On peut vérifier que le potentiel v satisfait l'équation:

$$\nabla^2 v = 4\pi G\rho\Delta$$

et, à chaque discontinuité, les conditions de continuité de V et de $(-4\pi G\rho W + dV/dr)$.

La force de gravité a pour composantes:

suivant le rayon:

$$\rho\frac{\partial v}{\partial r} + \rho\frac{\partial}{\partial r}\left(w\frac{dV_0}{dr}\right) - \rho\Delta\frac{dV_0}{dr}$$

et suivant le méridien:

$$\rho r^{-1}\left(\frac{\partial v}{\partial\theta} + \frac{dV_0}{dr}\frac{\partial w}{\partial\theta}\right),$$

d'où:

$$\mathscr{E}_g = (-1/10)\int_0^a \rho[\{dV/dr + d(W\,dV_0/dr)/dr - \delta\, dV_0/dr\} rW + \\ + 6U(V + W\,dV_0/dr)] r\, dr. \qquad (15)$$

En remplaçant dans les composantes de la force de gravité, les potentiels par leur expression, on trouve finalement l'énergie potentielle moyenne de gravitation:

$$\mathscr{E}_g = (-\rho_m/10) 4\pi G\rho_m I_g, \qquad (16)$$

avec:

$$5I_g = (\rho_c/\rho_m)\int_0^b [\{rFc' + (5\rho_c/3\rho_m)(rW - 12rU)\}rW + 6F_c rU]\,dr +$$

$$+ \int_b^a [\{rF_m' + (5/3)(rW - 12rU) - (20b^3/3r^3)(-1 + \rho_c/\rho_m)(rW - 3rU)\}rW +$$

$$+ 6F_m rU]\,dr.$$

Application numérique:

Les composantes du déplacement, pour la solution I, ont été calculées; si l'on prend: $\rho_c/\rho_m = 2{\cdot}4$, il vient:

$$F_c(r) = 12h_2(r) - 37{\cdot}245(r/a)^2$$

$$F_m(r) = 5g_2(r) + 35{\cdot}180(r/a)^2 - 36{\cdot}920(b/r)^3.$$

Par intégration graphique on trouve: $\rho_m I_g/I_0 = -0{\cdot}506$.
Si on prend: $\rho_m = 4{\cdot}5\,\mathrm{g/cm^3}$, il vient alors, d'après (14') et (16):

$$T = T_0[1 - G\rho_m T_0^2 \rho_m I_g/(\pi I_0)]^{-\frac{1}{2}} = 71/1{\cdot}37 = 51{\cdot}8\,\mathrm{min}.$$

(*b*) *Recherche d'une meilleure solution d'essai*

Si l'on tient compte de la gravitation, les équations du mouvement dans chaque milieu sont les équations (9); dans le noyau, ces équations ont la même forme (9') que dans une sphère homogène (dV_0/dr étant le même). Nous écrirons les équations (9') sous la forme:

$$Y = -\rho p^2 W$$

$$X = -\rho p^2 U.$$

Dans le manteau la solution exacte devrait satisfaire le système:

$$Y - (4\pi G\rho_m^2/3)(-1 + \rho_c/\rho_m)[d(Wb^3r^{-2})/dr - \delta b^3 r^{-2}] = -\rho p^2 W$$

$$X - (4\pi G\rho_m^2/3)(-1 + \rho_c/\rho_m)Wb^3r^{-2} = -\rho p^2 U.$$

Nous prendrons pour solution d'essai une solution des équations différentielles:

$$Y = -\rho p_0^2 W$$

$$X = -\rho p_0^2 U$$

dans le noyau et le manteau, avec une valeur d'essai p_0.

Dans la formulation (14) du principe de Rayleigh, on remplacera $\mathscr{T}_0$ et $\mathscr{E}_g$ par leurs expressions (7) et (15). L'expression de $\mathscr{E}_0$, (8), peut être transformée dans le cas d'un milieu homogène. Il vient alors:

$$\mathscr{E}_g + \mathscr{E}_0 = (1/10)\{-\int_0^a (WYr^2 + 6UXr^2)\,dr +$$

$$+ [(\lambda\delta + 2\mu W')Wr^2 + 6\mu rU(rU' - U + W)]_0^a -$$

$$- (4\pi G\rho_m^2/3)(-1 + \rho_c/\rho_m)\int_b^a b^3 r^2 W[\rho r^{-2} - (Wr^{-2})' - 6Ur^{-3}]\,dr\}.$$

Finalement:

$$(p^2-p_0^2)I_0 = H-(4\pi G\rho_m{}^2/3)(-1+\rho_c/\rho_m)J_g \qquad (17)$$

avec:

$$H = [(\lambda\delta+2\mu W')Wr^2+6\mu rU(rU'-U+W)]_0^a$$

$$J_g = 4\int_b^a (b/r)^3(rW-3rU)rW\,dr.$$

Le calcul de la solution d'essai est le même que pour une sphère homogène gravitante; dans le noyau, où $\mu = 0$, l'équation (12) devient:

$$\lambda_c\rho_c p_0{}^2x^2 = \rho_c p_0{}^2(\rho_c p_0{}^2+16\pi G\rho_c{}^2/3)-32\pi^2G^2\rho_c{}^2/3.$$

Posons:

$$q_c{}^2 = 4\pi G\rho_c/(3p_0{}^2).$$

La racine de l'équation est alors:

$$\alpha_c{}^2 = (\rho_c p_0{}^2/\lambda_c)(1+4q_c{}^2-6q_c{}^4).$$

Les composantes du déplacement ont pour amplitude:

$$\left.\begin{aligned} rU &= h_2-\eta_c h_1{}^0+C_c'(r/a)^2 \\ rW &= h_1{}^0-3h_2-2\eta_c h_1{}^0+2C_c'(r/a)^2 \end{aligned}\right\}$$

h_2 et $h_1{}^0$ sont, dans le cas où $\alpha_c{}^2$ est positif, les mêmes fonctions que dans l'expression (3) pour une sphère non gravitante, mais la variable est ici $\alpha_c r$ au lieu de $h_c r$. Dans le cas où $\alpha_c{}^2$ est négatif, les fonctions de Bessel $J_{n+\frac{1}{2}}(\alpha_c r)$ sont remplacées par:

$$I_{n+\frac{1}{2}}(|\alpha_c r|) = \exp[-(n+\tfrac{1}{2})i\pi/2]J_{n+\frac{1}{2}}(\alpha_c r).$$

On a:

$$\eta_c = q_c{}^2/(2q_c{}^2-1).$$

L'amplitude de la dilatation a pour expression:

$$\delta_c = -\alpha_c{}^2(1-2\eta_c)h_2$$

et celle du potentiel supplémentaire:

$$V_c = 4\pi G\rho_c[(1-2\eta_c)h_2+(2q_c{}^2-1)(C_c'/3q_c{}^2)(r/a)^2].$$

Si $\lambda = \mu$ dans le manteau, l'équation (12) peut s'écrire:

$$3X^2-4X(1+q_m{}^2)+1+4q_m{}^2-6q_m{}^4 = 0,$$

en posant:

$$q_m{}^2 = 4\pi G\rho_m/(3p_0{}^2), \qquad X = \mu x^2/(\rho_m p_0{}^2).$$

Soient $x^2 = \alpha_m{}^2$, et $x^2 = \beta_m{}^2$, les racines de cette équation. Les composantes du déplacement ont pour amplitude:

$$\left.\begin{aligned} rU &= g_2-(l_1{}^0-2l_2)+H(r)+D_m'(b/r)^3 \\ rW &= g_1{}^0-3g_2-6l_2+2H(r)-3D_m'(b/r)^3 \end{aligned}\right\}$$

avec:

$$H = \epsilon_m l_1-\eta_m g_1{}^0+C_m'(r/a)^2$$

g et l sont les mêmes fonctions que celles des expressions (6), mais les variables sont respectivement $\alpha_m r$ et $\beta_m r$ au lieu de $h_m r$ et $k_m r$.

$$\epsilon_m = (X_\beta^2 - 1)/(X_\beta^2 - 1 + 2q_m^2), \qquad \eta_m = q_m^2/(X_\alpha^2 - 1 + 2q_m^2).$$

La dilatation a pour amplitude:

$$\delta_m = -\alpha_m^2(1-2\eta_m)g_2 - 2\epsilon_m l_2 \beta_m^2$$

et le potentiel supplémentaire:

$$V_m = 4\pi G\rho_m[(1-2\eta_m)g_2 + 2\epsilon_m l_2 + (2-q_m^{-2})(C_m'/3)(r/a)^2 - \\ -(1+q_m^{-2}/3)D_m'(b/r)^3].$$

La solution exacte devrait satisfaire, en plus des équations du mouvement, 8 conditions aux limites, soit: les 5 conditions dejà énoncées en I.2, pour une sphère non gravitante, auxquelles il faut ajouter 3 conditions pour le potentiel:

$$r = b : V_m = V_e \quad \text{et} \quad (-4\pi G\rho W + dV/dr)_m = (-4\pi G\rho W + dV/dr)_e$$
$$r = a : dV/dr - 4\pi G\rho W = V(a)\, d(a^3 r^{-3})/dr.$$

Avec la solution d'essai adoptée, la valeur de p tirée de l'équation (17) est nécessairement trop grande. La meilleure valeur serait la plus petite valeur possible tirée de (17). En plus du paramètre p_0, I_0, H et J_g contiennent 8 constantes arbitraires. L'équation (17) pourrait donc donner lieu à un calcul variationnel, avec différentes possibilités suivant les conditions auxquelles on astreint la solution d'essai.

Les constantes sont entièrement déterminées (ainsi que p_0) si la solution d'essai satisfait toutes les conditions aux limites; on a alors $H = 0$, mais la solution choisie de la sorte n'est sans doute pas la meilleure. Nous avons choisi une solution d'essai satisfaisant toutes les conditions aux limites sauf une; en faisant varier p_0, on voit que la meilleure solution correspond à une valeur de p_0 voisine de celle qui annulle le deuxième membre de (17). La condition non satisfaite est ici l'annulation de la tension normale à la surface; il y a évidemment quatre autres possibilités, et l'essai d'une condition non satisfaite sur le noyau aurait peut-être été meilleur.

Application numérique:

Des essais pour différentes valeurs de p_0 ont montré que la compensation des deux termes du second membre de (17) se situe au voisinage de $T_0 = 58$ min. Nous avons donc calculé la solution correspondant à cette période (solution II):

$q_c^{-2} = 1{\cdot}0804$ d'où $|\alpha_c b| = 0{\cdot}4463$ (la racine de (13) est alors négative)

$\eta_c = 1{\cdot}0874,$

$q_m^{-2} = 2{\cdot}5928$ d'où $\alpha_m a = 1{\cdot}0808, \quad \beta_m a = 2{\cdot}1489$

$\eta_m = 2{\cdot}6707 \quad \epsilon_m = 0{\cdot}38087.$

Les 7 conditions aux limites conduisent aux valeurs:

$A_m' = -1{\cdot}3321 \quad B_m' = -2{\cdot}95545 \quad A_e' = 0{\cdot}11217 \quad C_m' = -1{\cdot}4726$

$\bar{A}_m' = -0{\cdot}034699 \quad \bar{B}_m' = -0{\cdot}11048 \quad C_e' = 1{\cdot}4437 \quad D_m' = 1.$

Dans le noyau, la partie principale du déplacement est donnée par les termes en $(r/a)^2$; les amplitudes des composantes du déplacement sont représentées sur la Figure 3.

Comme à la surface la tension normale n'est pas nulle, on a:

$$H = a\lambda_m(\delta + 2W')aW = 0{\cdot}8145\lambda_m/a.$$

Les intégrales J_g et I_0 ont été évaluées graphiquement. En portant les valeurs numériques dans (17), il vient:

$$(p/p_0)^2 - 1 = 0{\cdot}1313, \quad \text{soit: } T = 54{\cdot}6\,\text{min}$$

ce qui représente un résultat meilleur que celui obtenu plus haut: 51·8 min avec la solution I.

3. Application du principe de Rayleigh aux modèles (A) et (B) de Bullen

On utilise encore la relation (14) avec les expressions (7) et (8′) pour I_0 et $\mathscr{E}_0$; mais dans l'expression (15) de $\mathscr{E}_g$ on doit prendre:

$$dV_0/dr = -g,$$

intensité de la pesanteur à la distance r du centre

$$d^2V_0/dr^2 = -4\pi G\rho + 2g/r$$
$$V = (4\pi G/5)(2J_1 - 3J_2)$$
$$r\,dV/dr = (4\pi G/5)(5\rho rW - 6J_1 - 6J_2),$$

où:

$$J_1 = \frac{1}{r^3} \cdot \int_0^r \rho r_1^2(r_1W + 3r_1U)\,dr_1$$

$$J_2 = r^2 \int_r^a \rho(r_1W - 2r_1U)\,dr_1/r_1^3.$$

En fait seul le calcul de l'intégrale J_2 est nécessaire; en effet l'expression de $\mathscr{E}_g$ peut s'écrire:

$$\mathscr{E}_g = (-1/10)\int_0^a \rho[(rV' - 4\pi G\rho rW)rW - g(6U - 2W - \delta r + rW')rW + 6rUV]\,dr$$

or:

$$\int_0^a \rho(rV' \,.\, rW + 6rUV)\,dr = (1/4\pi G)\int_0^a (rV' + 3V)^2\,dr = 4\pi G\int_0^a (\rho rW - 3J_2)^2\,dr$$

car $\nabla^2 v = 4\pi G(w\rho' + \rho\Delta)$, et V satisfait la condition à la surface.

Finalement:

$$\mathscr{E}_g = (-2\pi G/5)\int_0^a [3J_2(3J_2 - 2\rho rW) + (g/\pi Gr)(rW - 3rU)\rho rW]\,dr.$$

$\mathscr{E}_g$ s'obtient par une double intégration.

Application numérique.

Modèle (A) de Bullen.

(a) *Solution d'essai I*

Les intégrales I_0 et $\mathscr{E}_0$ ont déjà été calculées dans la premiére partie. L'évaluation de $\mathscr{E}_g$ montre que:

$$\mathscr{E}_g/\mathscr{E}_0 = 0{\cdot}657$$

d'où une valeur approchée par défaut de la période:

$$T = 51{\cdot}5 \text{ min.}$$

L'effet des forces de gravitation est donc une diminution de la période de l'ordre de 20 pour cent, comme pour le modèle précédent.

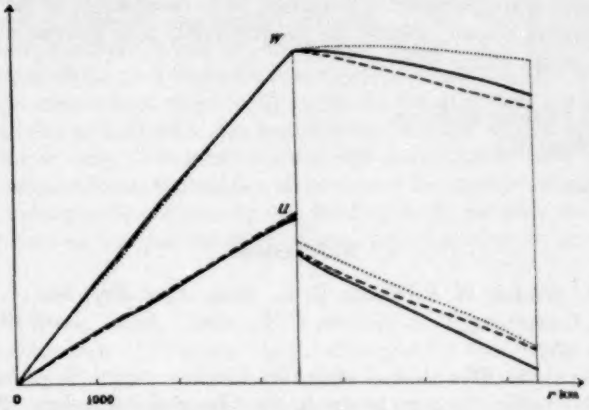


FIG. 3.—Variation des composantes du déplacement avec le rayon:
——— solution exacte pour le modèle (B) de Bullen (d'après Pekeris).
– – – – – solution d'essai I.
..... solution d'essai II.

(b) *Solution d'essai II.*

On trouve cette fois: $\mathscr{E}_g/\mathscr{E}_0 = 0{\cdot}562$ mais le résultat final est à peu près le même:

$$T = 51{\cdot}8 \text{ min.}^*$$

Ainsi la solution II, meilleure que la solution I pour une sphère à manteau et noyau homogènes, perd beaucoup de son intérêt pour un modèle hétérogène, la solution I étant beaucoup plus facile à calculer.

Modèle (B) de Bullen (1950)

La solution II a été essayée sur le modèle (B) de Bullen qui diffère du modèle (A) par une densité un peu plus grande dans le noyau, et un peu plus faible dans le manteau. On obtient le même résultat qu'avec le modèle (A).* Mais si on tient

* Par suite d'une erreur dans le calcul de l'énergie de gravitation, la valeur de la période publiée en 1957 était trop grande.

compte de la rigidité de la graîne, la période d'oscillation est diminuée assez sensiblement:

pour $\mu = 1{\cdot}5\ 10^{12}$ cgs, la période est diminuée de 1·3 min environ
pour $\mu = 3\ 10^{12}$ cgs, la période est diminuée de 2·8 min environ
pour $\mu = 3{\cdot}6\ 10^{12}$ cgs, la période est diminuée de 3·2 min environ.

La période trouvée avec $\mu = 0$ dans la graîne est comparable à celle obtenue par Pekeris par une méthode variationnelle (1958); l'erreur est de 4 pour cent par défaut, la période exacte pour le modèle (B) étant 53·7 min (Alterman, Jarosch, Pekeris 1959).

Sur la Figure 3 on peut comparer les déplacements pour les solutions I et II avec le déplacement exact pour le modèle (B). (Les solutions ont été réduites de façon à avoir la même valeur pour W sur le noyau.) Les deux solutions d'essai sont très voisines de la solution exacte dans le noyau; dans le manteau apparaissent des différences, principalement à la surface où la composante U est trop grande pour les solutions d'essai; comme elle est très faible pour la vraie solution, il en résulte une grande erreur relative.

Institut de Physique du Globe,
Paris V.:
1960 *August.*

Références

Alterman, Z., Jarosch, H. & Pekeris, C. L., 1959. *Proc. Roy. Soc. A*, **252**, 80–95.
Benioff, H., Gutenberg, B. & Richter, C. F., 1954. *Trans. Amer. Geophys. Un.*, **35**, 984–985.
Bullen, K. E., 1950. *Mon. Not. R. Astr. Soc. Geophys. Suppl.*, **6**, 50–59.
Bullen, K. E., 1953. *An Introduction to the Theory of Seismology.* (Cambridge: University Press, 2ème édition.)
Gilbert, F. & MacDonald, G. J. F., 1960. *J. Geophys. Res.*, **65**, 675–693.
Jeffreys, H., 1935. *Mon. Not. R. Astr. Soc. Geophys. Suppl.*, **3**, 253–261.
Jobert, N., 1957. *C.R. Acad. Sci (Paris)*, **244**, 921–922; **245**, 1941–1943.
Lamb, H., 1882. *Proc. London Math. Soc.*, **13**, 189–212.
Love, A. E. H., 1911. *Some Problems of Geodynamics.* (Cambridge: University Press.)
Love, A. E. H., 1944. *The Mathematical Theory of Elasticity*, 4ème édition. (New York: Dover Publications.)
Matumoto, T. & Satô, Y., 1954. *Bull. Earthq. Res. Inst.*, **32**, 247–258.
Pekeris, C. L. & Jarosch, H., 1958. *Contributions in Geophysics.* (Pergamon Press.)
Lord Rayleigh, 1894. *The Theory of Sound.* 2ème édition. (New York: Dover Publications).
Sezawa, K., 1929. *Bull. Earthq. Res. Inst.*, **6**, 1–18.
Stoneley, R., 1926. *Mon. Not. R. Astr. Soc. Geophys. Suppl.*, **1**, 356–359.
Stoneley, R., 1953. *Bull. Seismol. Soc. Amer.*, **43**, 127–135.
Watson, G. N., 1944. *Theory of Bessel Functions*, 2ème édition. (Cambridge: University Press.)

Torsional Oscillations of the Earth

H. Takeuchi

" Some say the Earth was fevrous and did shake."—*W. Shakespeare, Macbeth*, II, 3.

Summary

Making use of the variational calculus method developed in a previous paper, we calculate the free periods of torsional oscillations for two Earth models. In the Model 1, the Earth is assumed to be composed of homogeneous mantle and core. In the Model 2, we use the mantle structure model by Jeffreys and Bullen. In order to see the dependence of the periods on the core rigidity, we calculate the periods when the core is fluid or perfectly rigid. In the fluid core model, we work out the periods when the coupling between the mantle and core is perfect or zero. The present study will show that in order to get useful information on the rigidity of the core or the strength of coupling between the mantle and core by this kind of study, we must make the observations of free periods with an accuracy of several ten seconds.

1. Introduction

In a previous paper (Takeuchi 1959), subsequently referred to as Paper 1, using the variational calculus, we calculated the periods of torsional oscillations of the Earth, the internal constitution of which was inferred by Sir Harold Jeffreys and Bullen. In making the calculation, we assumed that the rigidity of the core is zero and due to a relative slip at the core boundary the core has nothing to do with the torsional oscillations. In a paper on the compressional oscillations of the Earth, Alterman, Jarosch & Pekeris (1959) attacked the same problem. According to them, the periods of the fundamental oscillation for the azimuthal wave number $n = 2$ and 4 are 44·1 and 21·9 min, respectively. These values agree with 43·4 and 21·5 min in Paper 1. In order to check the results obtained by her variational calculus method, Jobert (1959) solved the problem by a finite difference method and got results in good agreement with those of Paper 1. Gilbert & MacDonald (1960), attacking the problem by the homogeneous layer approximation, obtained results disagreeing with those in Paper 1. The period 45·5 min for $n = 2$ obtained by them, for example, differs from our 43·4 min referred to above. They attributed the disagreement to the inaccuracy of our variational calculus approximation. It seems to the present author that the disagreement is due to the inaccuracy of their approximation. In fact, an unpublished paper by Sato, Landisman and Ewing gives results preferring ours to theirs. One of the expected results by this kind of study is the estimation of core rigidity. In the present paper, using the variational calculus method of Paper 1, we shall study the problems relating to this possibility.

2. Variational method

As was described in Paper 1, the fundamental equation for our present problem is

$$\mu\left(\ddot{l}-\frac{n(n+1)}{\xi^2}\right)+\dot{\mu}\left(\dot{l}-\frac{2}{\xi}l\right)+\rho p^2a^2l=0 \tag{2.1}$$

and the boundary conditions to be satisfied are

$$\dot{l}-\frac{2}{\xi}l=0 \quad \text{at} \quad \xi=1 \tag{2.2}$$

and

$$\mu\left(\dot{l}-\frac{2}{\xi}l\right) \text{ and } l \text{ are continuous at } \xi=\xi_0, \text{ say.} \tag{2.3}$$

In (2.1), ρ and μ are density and rigidity, each being assumed to be a function of ξ, n the degree of spherical harmonic, p the frequency, $\exp(ipt)$ being understood throughout the equations. ξ is the distance r from the centre of the sphere divided by the radius of the sphere a. ξ_0 in (2.3) is ξ corresponding to some surface of discontinuity within the sphere, if one exists. $\cdot$ denotes the differentiation with respect to ξ. As the surface of discontinuity $\xi=\xi_0$, we take the core boundary

$$\xi=\frac{6370-2900}{6370}.$$

Denoting the rigidity immediately above and under the core boundary by μ_m and μ_c, respectively, we write (2.3) in the form

$$l=l_c, \qquad \mu_m\left(\dot{l}-\frac{2}{\xi}l\right)=\mu_c\left(\dot{l}_c-\frac{2}{\xi}l_c\right) \quad \text{at} \quad \xi=\xi_0, \tag{2.4}$$

l and l_c being the solution in the mantle $\xi\geqslant\xi_0$ and core $\xi\leqslant\xi_0$ respectively. Rearranging (2.4) in the following form will eliminate an integration constant in l_c.

$$\frac{\xi\dot{l}-2l}{l}=A, \qquad A=\frac{\mu_c}{\mu_m}\left(\frac{\xi\dot{l}_c-2l_c}{l_c}\right)\xi=\xi_0. \tag{2.5}$$

Since the boundary conditions (2.2) and (2.5) contain no n, we may take the following trial function in the variational calculation for any n.

$$l=l(m)=\xi^m(1+\alpha\xi+\beta\xi^2), \tag{2.6}$$

where m is a parameter and α and β are constants to be determined by (2.2) and (2.5) as follows:

$$\begin{aligned}(m-1)\alpha+m\beta &= -(m-2),\\ \{(m-1)-A\}\alpha\xi_0+(m-A)\beta\xi_0^2 &= A-(m-2).\end{aligned} \tag{2.7}$$

For $m=2$, 6 and 20, for example, we have

$$\beta=\frac{1{\cdot}83573487A}{1{\cdot}45525903A-0{\cdot}91051805}, \qquad \alpha=-2\beta, \tag{2.8}$$

$$\beta=\frac{5{\cdot}17867436A-16{\cdot}71469740}{3{\cdot}27629513A-13{\cdot}65777080}, \qquad 5\alpha=-4-6\beta \tag{2.9}$$

$$\beta = \frac{9{\cdot}19466248A - 155{\cdot}69858713}{5{\cdot}25670763A - 94{\cdot}23933321}, \qquad 19\alpha = -18 - 20\beta \tag{2.10}$$

respectively.

As in Paper 1, the variational equation equivalent to (2.1) to (2.3) is

$$\int_0^1 \rho p^2 a^2 f(\rho)\, d\xi - \int_0^1 \mu\{Nf(\mu, N) + f(\mu, 1)\}\, d\xi = 0;$$

$$N = n(n+1), \qquad f(\rho) = \xi^2, \qquad f(\mu, N) = l\delta l,$$

$$f(\mu, 1) = \dot{l}\delta\dot{l} - \frac{2}{\xi}(\dot{l}\delta l + l\delta\dot{l}) + \frac{2}{\xi^2} l\delta l. \tag{2.11}$$

Putting

$$l = \sum_m A_m l(m), \tag{2.12}$$

and making variation of undetermined constants A_m, we have simultaneous linear equations among A_m. The coefficient of $A_{\tilde{m}}$, say, in the mth equation is obtained by replacing l and δl in the left-hand side of (2.11) by $l(\tilde{m})$ and $l(m)$, respectively. Equating the determinant formed by the coefficients to zero, we have approximate eigenvalues

$$P = \frac{pa}{V_{s.0}}, \qquad V_{s.0} = \sqrt{\frac{\mu_0}{\rho_0}}, \tag{2.13}$$

where ρ_0, μ_0 and $V_{s.0}$ are density, rigidity and shear wave velocity at the surface of the Earth $\xi = 1$.

3. Fluid core

When the rigidity of the core μ_c is zero, we have $A \sim \xi\dot{l} - 2l = 0$ and by (2.8) to (2.10)

$$\begin{aligned} l(2) &= \xi^2, \\ l(6) &= \xi^6(1 - 2{\cdot}26858790\xi + 1{\cdot}22382325\xi^2), \\ l(20) &= \xi^{20}(1 - 2{\cdot}68648567\xi + 1{\cdot}65216138\xi^2) \end{aligned} \tag{3.1}$$

in the mantle. When there is a slip at the core boundary and the existence of the core has nothing to do with the torsional oscillations, we have $l_c(\xi) \equiv 0$ and we carry out the integration (2.11) only in the interval $\xi_0 \leqslant \xi \leqslant 1$. As in Paper 1, we take two Earth models. In Model 1, we have $\rho = \text{const} = \rho_0$ and $\mu = \text{const} = \mu_0$ in the mantle, and in Model 2, we take the mantle structure inferred by Sir Harold Jeffreys and Bullen. The following power series approximations for ρ, λ and μ are used in the calculations.

0 to 400 km depth

$$\rho = 8{\cdot}60500000 - 5{\cdot}30833333\xi, \qquad \lambda = 7{\cdot}55100000 - 6{\cdot}86544444\xi,$$
$$\mu = 5{\cdot}26800000 - 4{\cdot}67133333\xi, \tag{3.2}$$

400 to 1 000 km depth

$$\rho = -48{\cdot}25925000 + 129{\cdot}31100000\xi - 78{\cdot}89952777\xi^2,$$
$$\lambda = -44{\cdot}99163332 + 118{\cdot}21304438\xi - 73{\cdot}63955924\xi^2, \qquad (3{\cdot}3)$$
$$\mu = -35{\cdot}46730000 + 93{\cdot}72393329\xi - 58{\cdot}61107777\xi^2,$$

1 000 to 2 900 km depth

$$\rho = 7{\cdot}51416666 - 3{\cdot}36194444\xi, \qquad \lambda = 8{\cdot}71433333 - 7{\cdot}57322222\xi, \qquad (3{\cdot}4)$$
$$\mu = 5{\cdot}20150000 - 3{\cdot}92816666\xi,$$

the units of ρ, λ and μ being g/cm^3 and 10^{12} dyn/cm^2, respectively. The third approximations for P thus obtained coincide with those in Paper 1, and are shown in Table 1.

Table 1

	P	
n	Model 1	Model 2
2	2·412	3·606
3	3·790	5·579
4	5·045	7·292
8	9·624	12·94

It may be that the rigidity of the core is zero but the coupling between the mantle and core is perfect. In such a case, we have $l = l_c$ at the core boundary and we must carry out the ρ integration in (2.11) in the interval $0 \leqslant \xi \leqslant \xi_0$ also. This means that the core has no effect in the restoring force but works as an additional mass in the torsional oscillations. As trial functions for $l_c(\xi)$, we take (3.1), which satisfy the boundary conditions $l = l_c$ at $\xi = \xi_0$.

In carrying out the calculations for Models 1 and 2, we assume that the density in the core ρ_c is constant and $\rho_c/\rho_m = 2$, where ρ_m is the density immediately above the core boundary. In Model 1, in which $\rho = \text{const} = \rho_0 = \rho_m$ in the mantle, ρ_c/ρ_m is equal to ρ_c/ρ_0. The third approximations for P thus obtained are shown in Table 2.

Table 2

	P	
n	Model 1	Model 2
2	2·307	3·418
3	3·643	5·347
4	4·880	7·073
8	9·514	12·89

4. Perfectly rigid core

If the core were perfectly rigid, we would have $l_c = 0$ and $A = \infty$. By using (2.8) to (2.10) we get $l(m)$ as follows.

$$l(2) = \xi^2(1 - 2{\cdot}52289776\xi + 1{\cdot}26144888\xi^2),$$
$$l(6) = \xi^6(1 - 2{\cdot}69677943\xi + 1{\cdot}58064953\xi^2),$$
$$l(20) = \xi^{20}(1 - 2{\cdot}78855739\xi + 1{\cdot}74912952\xi^2). \qquad (4{\cdot}1)$$

The third approximations for P in this case are shown in Table 3. P's for Model 1 have already been obtained by Matsumoto & Sato (1954). Their results are

$$P = 2{\cdot}00,\ 3{\cdot}05,\ 4{\cdot}30,\ 5{\cdot}40 \quad \text{for} \quad n = 1,\ 2,\ 3,\ 4. \tag{4.2}$$

Our P's for $n = 1$ and 2 agree with theirs well, but for $n = 3$ and 4, our P's are significantly smaller than theirs. Since the variational calculus method gives P larger than the exact value, and since our method gives P for different n in a systematic way, the disagreement is probably due to some mistakes in their calculations.

5. Homogeneous core

In order to attack the problem when μ_c is neither 0 nor infinite, we assume $\rho = \text{const} = \rho_c$, $\mu = \text{const} = \mu_c$, and shear wave velocity $= \text{const} = V_{s.c} = \sqrt{(\mu_c/\rho_c)}$ in the core. The solution l_c of (2.1) in this case is

$$l_c \sim \chi^{\frac{1}{2}} J_{n+\frac{1}{2}}(\chi), \qquad \chi = \frac{pa\xi}{V_{s.c}}, \tag{5.1}$$

and for (5.1) we get A in (2.5) as follows:

$$A = \frac{\mu_c}{\mu_m}\left\{\chi_0 \frac{J_{n-\frac{1}{2}}(\chi_0)}{J_{n+\frac{1}{2}}(\chi_0)} - (n+2)\right\},$$
$$\chi_0 = \frac{pa\xi_0}{V_{s.c}}. \tag{5.2}$$

Practical numerical computations are carried out as follows. By using (2.8) to (2.10), we calculate $l(m)$ for a chosen A, and putting $l(m)$ thus obtained into the left-hand side of (2.11), we get the coefficients in the characteristic determinant. The integrations in (2.11) are carried out only in the interval $\xi_0 \leqslant \xi \leqslant 1$. The reason why we can omit the integrations in the interval $0 \leqslant \xi \leqslant \xi_0$ is that since (5.1) is the exact solution of (2.1), the ρ and μ integrals in (2.11) cancel with each other. Equating the determinant thus obtained to zero, we get $P = pa/V_{s.0}$. On the other hand, solving (5.2) for the given A, we get

$$\chi_0 = \frac{pa\xi_0}{V_{s.c}} = \frac{\xi_0 V_{s.0}}{V_{s.c}} P. \tag{5.3}$$

Since ξ_0, $V_{s.0}$ and P are known in the right-hand side of (5.3), we can now determine $V_{s.c}$ which gives P obtained before.

The right-hand side of (5.2) becomes zero for a certain χ_0. Thus

$$\text{for } n = 2,\ 3 \text{ and } 4,\ \chi_0 = 2{\cdot}501,\ 3{\cdot}864 \text{ and } 5{\cdot}094 \tag{5.4}$$

will make A in (5.2) vanish. As was said immediately after (5.2), we have only to carry out the integrations in (2.11) in the interval $\xi_0 \leqslant \xi \leqslant 1$. Thus P's in this case are nothing but those in Table 1. Table 1, (5.4) and (5.3) will determine $V_{s.c}/V_{s.0}$ giving P's in Table 1. Thus for the Earth Model 1,

$$V_{s.c}/V_{s.0} = 0{\cdot}5254,\ 0{\cdot}5343 \text{ and } 0{\cdot}5395 \quad \text{for} \quad n = 2,\ 3 \text{ and } 4 \tag{5.5}$$

and for the Earth Model 2,

$$V_{s.c}/V_{s.0} = 0{\cdot}7854,\ 0{\cdot}7865 \text{ and } 0{\cdot}7798 \quad \text{for} \quad n = 2,\ 3 \text{ and } 4 \tag{5.6}$$

will give P's in Table 1.

6. Discussions

The following are numerical data for the Earth Model 2

$$\rho_0 = 3{\cdot}3\,\mathrm{g/cm^3}, \quad \rho_m = 5{\cdot}683\,\mathrm{g/cm^3}, \quad \rho_c = 11{\cdot}366\,\mathrm{g/cm^3},$$
$$\mu_0 = 0{\cdot}6\times 10^{12}\,\mathrm{dyn/cm^2}, \quad V_{s.0} = 4{\cdot}26\,\mathrm{km/s} \tag{6.1}$$

and Tables 1, 2 and 3 give the periods of torsional oscillations shown in Table 4. When the coupling between the mantle and core is complete and there is no slip at the core boundary, the periods in the columns under Table 2, Table 3 and

Table 3

P

n	Model 1	Model 2
1	1·998	3·690
2	3·040	4·887
3	4·132	6·230
4	5·236	7·617
8	9·633	12·95

Table 4

n	Table 2	Table 1	Table 3
1			42·4 min
2	45·8 min	43·4 min ($\mu = 1{\cdot}275\times 10^{12}$)	32·0
3	29·3	28·1 ($\mu = 1{\cdot}278\times 10^{12}$)	25·1
4	22·1	21·5 ($\mu = 1{\cdot}257\times 10^{12}$)	20·6
8	12·2	12·1	12·1

Table 1 in Table 4 are those for $\mu_c = 0$, $\mu_c = \infty$ and μ_c in the parentheses. The periods in the column under Table 1 in Table 4 are also those for the fluid core when due to a slip at the core boundary the core has nothing to do with the torsional oscillations of the Earth. According to Table 4, in order to make a difference between μ_c and $\mu_c = 1{\cdot}275\times 10^{12}\,\mathrm{dyn/cm^2}$ by observing the torsional oscillation of degree 2, for example, we must make a difference between the period 45·8 min and 43·4 min. The same accuracy in the period determination will be necessary in order to get useful information on the coupling between the mantle and the fluid core. For the oscillations of the larger wave number n, we need more accurate determinations of the periods. In short, in order to get useful information on the rigidity of the core or the strength of coupling between the mantle and core by this kind of study, we must make the observations of periods with the accuracy of several tens of seconds.

Department of Geophysics,
University of Tokyo,
Japan:

1960 *August.*

References

Alterman, Z., Jarosch, H. & Pekeris, C. L., 1959. *Proc. Roy. Soc. A*, **232**, 80–95.
Gilbert, F. & MacDonald, G. J. F., 1960. *J. Geophys. Res.*, **65**, 675–693.
Jobert, N., 1959. *C.R. Acad. Sci. (Paris)*, **249**, 1014–1016.
Matsumoto, T. & Sato, Y., 1954. *Bull. Earthq. Res. Inst.*, **32**, 247–258.
Takeuchi, H., 1959. *Geophys. J.*, **2**, 89–100.

V. Thermal State

J. A. JACOBS, Some aspects of the thermal history of the Earth 267

J. VERHOOGEN, Heat balance of the Earth's core 276

E. C. BULLARD & A. DAY, The flow of heat through the floor of the Atlantic Ocean 282

Some Aspects of the Thermal History of the Earth

J. A. Jacobs

" 'Where do you come from?' said the Red Queen, 'and where are you going?' "—
Lewis Carroll, Through the Looking Glass.

Summary

A number of thermal problems connected with the Earth are discussed—the early thermal history of the Earth and some of its consequences, the formation and state of the inner core, convection in the core and its relation to the Earth's magnetic field, and conditions in the upper mantle.

1. Introduction

Of all the problems connected with the interior of the Earth, that of the variation of its temperature, both with time and with depth, remains one of the most difficult to settle. Verhoogen (1956) has given a table which lists various estimates that have been made between 1915–1955 of the present temperature at depths of 100 km, 2 900 km and at the centre of the Earth. To this table may now be added some six or more further estimates. It is instructive, indeed, to plot the "secular variation" of the Earth's temperature over the last 45 years!

The temperature distribution is the key to a number of other problems that are currently engaging the attention of many geophysicists. The question of possible phase changes in the crust and upper mantle (MacDonald & Ness 1960) and any attempt to discuss the physics of continental drift are but two examples. Rather than attempt to estimate the actual temperature distribution, some aspects of certain geothermal problems are discussed in the following sections.

2. The early thermal history of the Earth

All thermal histories of the Earth are based on either a hot or on a cold origin; should a cold origin lead at some later stage in the evolution of the Earth to a molten state, the subsequent thermal histories, whatever the initial origin, would be the same. A cold origin is usually preferred now, and quite low initial temperatures (well below the melting point of the silicate rocks that form the mantle) are generally postulated. Ringwood (1960) however has given rather convincing reasons for believing that an Earth formed by accretion would at an early stage in its development become molton. A molten Earth also provides by fractional crystallization, the only reasonable explanation of the concentration of U and Th in its outer regions.

Estimates of the time of accretion are of the order of 10^8 years. Possible heat sources that could raise the temperature of the Earth during this period of accretion are the radioactive decay of both long-lived and short-lived isotopes, chemical

reactions, and the conversion of kinetic energy into thermal energy. Because of the comparatively short time of accretion, the temperature increase due to the radioactive decay of long-lived isotopes is small, of the order of 150° C (MacDonald 1959). Short-lived radioactive isotopes could have contributed to the initial heat of the Earth if the time between the formation of the elements and the aggregation of the Earth was short compared with the half lives of the isotopes. The important short-lived isotopes are U^{236}, Sm^{146}, Po^{244} and Cm^{247}, all of which have half lives sufficiently long to have heated up the Earth during the 10^7–10^8 years after the initial formation. MacDonald (1959) estimates that if all this heat was retained by the Earth, a temperature increase of the order of 2000–3000° C may be possible.

The temperature of the material within the aggregating Earth will also increase because of adiabatic compression. Although data (particularly on the variation of the coefficient of thermal expansion) are rather uncertain, a rise in temperature from this source of several hundred degrees seems likely. However the largest source of available energy is the potential energy due to the mutual gravitational attraction of the particles of the dust cloud. This energy, upon aggregation, is either converted into internal energy or radiated away. It is difficult to estimate the total contribution from this source because of the uncertainty of the physical process of accretion. The result depends quite critically on the temperature attained at the surface of the aggregating Earth and on the transparency of the surrounding atmosphere to radiation.

Comparatively low surface temperatures (of the order of a few hundred degrees) have been predicted, mainly because the atmosphere of the primitive Earth was assumed transparent so that the thermal energy of the impinging particles was immediately re-radiated into space. Ringwood (1960) has shown, however, that during these early years, the primitive Earth will have a large reducing atmosphere. In the presence of these reducing agents (chiefly carbon and methane), the accreting material will be reduced to metallic alloys—principally of iron, nickel and silicon. The outer regions of the Earth will thus be metal rich and dense (referred to zero pressure) compared to the interior. Such a state is gravitationally unstable, and convective overturn will follow leading to a sinking of the metal rich outer region into the centre. This would release further heat due to the energy of gravitational rearrangement. Ringwood believes the whole process is likely to be catastrophic, since the overturn will be accelerated as the initial temperature rises.

Ringwood's (1960) theory of the early history of the Earth implies the production of an enormous atmosphere of CO_2, CO and H_2O due to reduction processes on the primitive Earth. A critical requirement of his theory is the dissipation of this atmosphere at an early stage, and in his paper he discusses factors which may have led to this. It is instructive to consider the effect on the Earth of the removal of the surface pressure of this atmosphere. It would lead to an expansion of the Earth and possible polymorphic phase changes in its interior. Suppose an expansion of 6 km in the radius r of the Earth took place, i.e. $(dr/r) \simeq 10^{-3}$. Then, assuming hydrostatic equilibrium, the relief of pressure dp necessary for such an expansion, is given by $dp = k(dV/V)$, where k is the incompressibility and $(dV/V) \simeq (3dr/r)$ is the volume change. Taking an average value of k for the Earth (4×10^{12} dyn/cm^{-2}), this leads to a value of $dp \simeq 10^{10}$ dyn/cm$^2 \simeq 10^4$ atmospheres.

Ringwood (personal communication) estimates the mass of this primitive atmosphere to be approximately $\frac{1}{3}$ that of the Earth, and the density at its base to

be of the order of 2 g/cm³. This would imply a scale height of approximately 1 000 km and a pressure at the Earth's surface of approximately 2.10^5 atmospheres. Allowing for the increase of k with pressure, the removal of such a pressure would lead to an expansion in the radius of the Earth of approximately 170 km. Thus an expansion of 100 km (due to the removal of an atmosphere with a pressure of approximately 10^5 atmospheres at the Earth's surface) in the early history of the Earth seems reasonable. Such an expansion may well have been sufficient to rupture the Earth's surface, and these initial "scars" may be the site of the mid-ocean ridges and continents, which developed differently throughout geologic time (cf. Wilson 1960). These surface expressions of the relief of pressure may be accompanied by polymorphic phase changes occurring deeper within the Earth (a pressure of 10^5 atmospheres is reached at a depth of approximately 300 km).

3. The Earth's inner core

In the last few years a number of estimates have been made of the variation with depth of the melting point of the materials that compose the Earth. Using solid-state theory and seismic data, Uffen (1952b) estimated the melting-point depth curve in the mantle, while Simon (1953) used a semi-empirical equation to estimate the melting point of iron and hence obtained melting-point depth curves in the core. Bullard (1954) has also used Simon's equation to estimate melting points in both the mantle and the core. A feature of all these estimates is much lower values for the melting point in the core than in the mantle. Recent experimental work by Strong (1959) on the fusion of iron up to a pressure of 96 000 atmospheres, also leads by extrapolation to comparatively low values for the melting point in the core.

Zharkov (1959) has estimated the melting point of the mantle on the theory that fusion is obtained when the density of thermal defects in the crystalline solid reaches a certain critical value. At depths greater than 1 000 km, his results are about 1 000° C lower than those obtained by Uffen (1952b). On the other hand, his estimates of the melting point of iron at high pressures are greater than those obtained by Simon (1953), his value of the melting point of iron at the core–mantle boundary being approximately the same as that of the mantle at that depth. He thus finds an almost smooth variation of the melting point–depth curve throughout the entire Earth. Gilvarry (1956) under certain assumptions, has given a theoretical justification for Simon's equation, and obtained fusion temperatures for both the mantle and the core. Unlike other results, he obtains melting points for the inner core in excess of those in the mantle (1957).

On the basis of Gilvarry's curves, the explanation put forward by Jacobs (1953) for the formation of a solid inner core as the Earth cooled from a completely molten state, still holds. A solid inner core and solid mantle would form if the fusion temperature for the inner core is higher than that for the mantle at its base (see Figure 1). In this case solidification would commence at the centre of the Earth. A solid inner core would continue to grow until a curve representing the adiabatic temperature intersected the melting point curve twice, once at A, the boundary between the core and mantle, and again at B. As the Earth continued to cool, the mantle would solidify from the bottom upwards, trapping a liquid layer between A and B. The mantle would cool at a relatively rapid rate, leaving this liquid layer essentially at its original temperature, insulated above by a rapidly thickening shell of silicates and below by the already solid (iron) inner core.

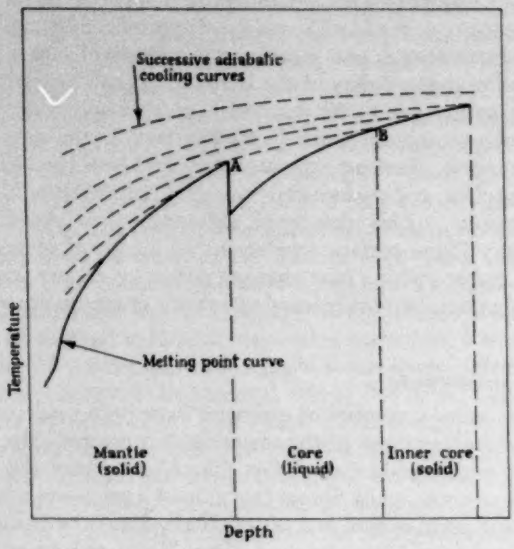


FIG. 1.
Possible formation of the Earth's inner core (after Jacobs 1953).

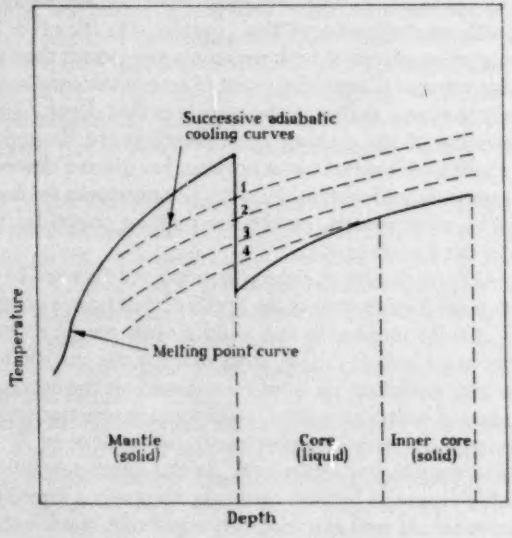


FIG. 2.
Alternative formation of the Earth's inner core.

In view of the other evidence, particularly the experimental work of Strong (1959), it is instructive to consider in some detail the (more probable) case when the melting point–depth curve in the core never rises above its value at the bottom of the mantle A. In this case as the Earth cooled, the entire core would be left liquid—at least initially. The critical point is whether the Earth could cool sufficiently for the temperature to fall below the melting point curve of the inner core and so finally leave it solid (see Figure 2). Two lines of evidence show that this is a possibility. Ringwood (1960) has pointed out that it is possible that the crystal mush at the bottom of the mantle may undergo an independent convection, whilst the overlying liquid is still crystallizing and convecting. This would greatly reduce the temperature throughout the solidified zone. Thus the actual temperature would quickly follow the curves $1 \rightarrow 2 \rightarrow 3$. When the mantle had completely solidified, the only mechanism for further cooling that has been considered until quite recently, is conduction—which is an extremely slow process. However, following a suggestion of Preston (1956) on the possibility of radiative heat transfer in the mantle, Clark (1957), Lubimova (1958) and Lawson & Jamieson (1958) have examined the possibility in some detail. Although differing in orders of magnitude, all these workers agree that the effect of radiative transfer is to increase considerably the effective thermal conductivity of the mantle. Lawson & Jamieson in particular believe that in the lower part of the mantle, the effective conductivity may be increased by a factor of a hundred. Moreover, although the ordinary molecular conductivity will probably decrease in the outer layers of the Earth due to the increase in temperature, Uffen (1952a) has shown that in the deeper parts of the mantle, the effect of pressure may dominate that of temperature and cause a considerable increase in the thermal conductivity. Finally Lubimova (1960) has considered one further process of heat transfer in the Earth's mantle—energy transfer by the excited states of atoms. At high temperatures, the contribution from the higher excitation energies increases (in normal conditions practically only the lowest excitation energy is important), and thus the rate of the excitation thermal conductivity must increase. Thus all the recent evidence is for a considerable increase in the effective conductivity with depth, and had the Earth cooled in its early history to curve 3 (Figure 2), it is more than possible that it could have cooled further throughout geologic time to curve 4, thus forming a solid inner core. The inner core has thus been slowly growing.

The gradient of the fusion curve for iron at high pressures may be rather flat, and it might therefore appear that if the further cooling suggested above actually took place, the whole core might solidify—which, of course, is contrary to seismic evidence. However present indications are that the outer core is likely to contain a fair amount of silicon (of the order of 20 per cent) which would lower its melting point compared to that of pure iron, thereby steepening the melting point gradient across the outer core. Any tendency of the Earth to heat up at depth due to long-lived radioactive isotopes may be offset by the increased values of the effective thermal conductivity. In any case most of the radioactive elements are concentrated in the crust and upper mantle, and Gast (1960) has shown that the abundance of potassium in the Earth may well be much lower than hitherto supposed. Taking the older (higher) abundances of the radioactive isotopes and assuming that there is no increase in the thermal conductivity with depth, Jacobs & Allan (1956) showed that the increase in temperature at the core–mantle boundary throughout geologic time was only about 300° C. This point will be considered again in the next section.

4. The Earth's magnetic field and convection in the core

The only satisfactory explanation of the Earth's main magnetic field is that it arises from a self acting dynamo action set up by motions in the fluid electrically conducting core. The nature and cause of these motions are unknown, although it has been shown that thermal convection provides the only satisfactory mechanism for their generation. Thermal convection will occur in the core if the transport of heat radially outwards exceeds the heat transport by thermal conduction alone. The material of the liquid (outer) core should be very nearly in a state of chemical equilibrium, uniformity being maintained by the mixing action of the convective motion. Thus it is unlikely that any motion is due to variations in the physical properties of the liquid outer core—rather it must be determined by boundary conditions. Three possible models have been proposed by Elsasser (1950). The first is one in which heat supplied by the inner core to the outer liquid layer exceeds the amount that can be carried away from there by conduction alone. The second is one in which the heat flow in the mantle adjacent to the core exceeds the purely conductive heat transport in the core itself. The third is a compromise model in which the heat supplied by the inner core is carried to the core–mantle boundary by convection and is then carried away by the mantle.

In the past it has been the first model that has been considered the most plausible, the heat source being supplied by the radioactive content of the inner core. It has been shown (Jacobs 1954) that very little radioactivity is needed—probably less than one per cent of the concentration in the Earth's crust. But even this small amount presents difficulties, since there are geochemical problems in having any radioactivity in the inner core and neutron activation analyses have failed to find any in iron meteorites (Reed & Turkevitch 1956; Bate & others 1958).

It is instructive therefore to consider Elsasser's second model, particularly in view of the large increase in the estimate of the effective thermal conductivity with depth in the mantle due to radiative heat transfer. If the heat carried away in the mantle exceeds that supplied by conduction alone in the core, the second model will hold. Extra heat could be supplied in the mantle by radioactive material distributed throughout the mantle, although such heat sources are not necessary. The convective process will occur if the heat conduction in the mantle is large enough, the net result being an average loss of heat from the core over geologic time. It is extremely difficult to estimate the relative magnitudes of the heat carried away by conduction in the mantle to that supplied by conduction in the core. The temperature gradient in the liquid core will be approximately the adiabatic, whilst that in the adjacent mantle should not be too different. Thus if the (effective) thermal conductivity of the mantle at the core–mantle boundary is greater than that of the core there, this model will hold. This may well be the case since the effective thermal conductivity of the mantle at depth may be as much as a hundred times that of surface rocks because of radiative transfer and thus greater than that in the metal rich core where the conductivity will be essentially molecular.

5. The upper mantle

In recent years there has accumulated a growing body of data which indicates that there are very fundamental differences in the constitution of the upper mantle beneath the oceans and beneath the continents. Jacobs (1960) has drawn attention

to the implications of the evidence of the increased number of heat flow measurements. The rough equality of the heat flow through all parts of the Earth, together with the fact that the oceanic crust, being relatively thin and composed mainly of basaltic materials, can produce at most 10 per cent of the observed values, whereas the whole of the heat flow could be produced by the radioactive isotopes in about 25 km of granite, emphasizes the striking difference between the oceanic and continental mantle. This difference is borne out by the results of a recent study by Dorman, Ewing & Oliver (1960) of shear velocity distributions in the upper mantle by mantle Rayleigh waves. Two hundred and fifty-six Rayleigh wave phase solutions were computed for 11 models of the continental and oceanic crust–mantle systems some with as many as 50 layers. The authors find a systematic difference in the sub-oceanic and sub-continental mantle down to a depth of 400 km. They also confirm the existence of a low-velocity layer, which is however shallower (beginning at about half the depth) under the oceans than under the continents. This difference in the structure of the upper mantle has also been reported by Miss Lehmann (1960) from an analysis of the travel times of seismic *P* and *S* waves. Such differences (down to about 400 km) must be considered in any theory of continental drift, which would imply that the continents drag the upper few hundred km of the mantle along with them, or fundamentally charge the mantle as they drift. These fundamental differences between sub-oceanic and sub-continental upper mantle are in accord with the idea that the whole of the crust of the Earth has been derived from lavas extruded from the mantle during geologic time (Wilson 1957, 1959). On this theory, it is not surprising that there are differences in upper mantle structure, in view of the very different crustal structure under continents and oceans. Recent studies (Russell 1956) of the isotope ratios in lead ores can be interpreted to mean that the majority of these ores have come directly from the mantle, giving further support to Wilson's thesis.

Lubimova (1958) has shown that the thermal conductivity of the upper layers of the Earth is determined almost entirely by its molecular component and that it has a minimum value (at a depth of 50–100 km) some 1·5–2·0 times lower than in normal surface rocks. The depth to the low velocity layer has been estimated at about 50 km under the oceans and 120 km under the continents. It is not unlikely that these two observations are related. Moreover because of the contribution from radiative transfer, the temperature gradient required to remove heat from the Earth's interior decreases as the effective thermal conductivity increases with temperature. As a result melting temperatures are most likely to be approached in the outer few hundred kilometres of the Earth. If the radioactivity is concentrated in the upper mantle, MacDonald (1959) estimates that the melting point may be approached, or even exceeded in the range 200–600 km below the surface. It is difficult to synthesize all these ideas and other possible relevant hypotheses (such as the possibility that the Earth is expanding and not contracting, a hypothesis that has drawn three independent letters to *Nature* within three months early in 1960) into a coherent whole—and even more difficult to distinguish between cause and effect in the different phenomena. It is all too easy to concentrate (in ignorance) on one point to the exclusion of another, while the first phenomenon may turn out to be but quite a secondary result of the more dominating role played by the other.

Acknowledgments

The author is grateful to Professor J. C. Jaeger, Australian National University, Canberra, and Professor K. E. Bullen, University of Sydney, who made it possible for him to spend two months in Australia, where most of this paper was written. Thanks are also due to Dr A. E. Ringwood for many stimulating discussions.

Geophysics Laboratory,
University of British Columbia,
Vancouver, Canada:
1960 *September.*

References

Bate, G. L., Potratz, H. A. & Huizenga, J. R., 1958. Thorium in iron meteorites. *Geochim. et Cosmochim. Acta* **14**, 118–125.

Bullard, E. C., 1954. *The Earth as a Planet* (G. P. Kuiper, ed.), Chapter 3. (Univ. Chicago Press, Chicago.)

Clark, S. P., Jr., 1957. Radiative transfer in the Earth's mantle. *Trans. Amer. Geophys. Un.*, **38**, 931–938.

Dorman, J., Ewing, M. & Oliver, J., 1960. Study of shear-velocity distribution in the upper mantle by mantle Rayleigh waves. *Bull. Seismol. Soc. Amer.*, **50**, 87–115.

Elsasser, W. M., 1950. Causes of motions in the Earth's core. *Trans. Amer. Geophys. Un.*, **31**, 454–462.

Gast, P. W., 1960. Limitations on the composition of the upper mantle. *J. Geophys. Res.*, **65**, 1287–1297.

Gilvarry, J. J., 1956. Equation of the fusion curve. *Phys. Rev.*, **102**, 325–331.

Gilvarry, J. J., 1957. Temperatures in the Earth's interior. *J. Atmos. Terr. Phys.*, **1**., 84–95.

Jacobs, J. A., 1953. The Earth's inner core. *Nature* (*Lond.*), **172**, 297.

Jacobs, J. A., 1954. Problems connected with the cooling of the Earth. *Trans. Amer. Geophys. Union*, **35**, 161–163.

Jacobs, J. A., 1960. Continental Drift. *Nature* (*Lond.*), **185**, 231–232.

Jacobs, J. A. & Allan, D. W., 1956. The thermal history of the Earth. *Nature* (*Lond.*), **177**, 155–157.

Lawson, A. W. & Jamieson, J. C., 1958. Energy transfer in the Earth's mantle. *J. Geol.*, **66**, 540–551.

Lehmann, I., 1960. Structure of the upper mantle as derived from the travel times of seismic *P* and *S* waves. *Nature* (*Lond.*), **186**, 956.

Lubimova, H. A., 1958. Thermal history of the Earth with consideration of the variable thermal conductivity of its mantle. *Geophys. J.*, **1**, 115–134.

Lubimova, H. A., 1960. On the process of heat transfer in the Earth's mantle. I.U.G.G. XIIth General Assembly, Helsinki.

MacDonald, G. J. F., 1959. Calculations on the thermal history of the Earth. *J. Geophys. Res.*, **64**, 1967–2000.

MacDonald, G. J. F. & Ness, N. F., 1960. Stability of phase transitions within the Earth. *J. Geophys. Res.*, **65**, 2173–2190.

Preston, F. W., 1956. Thermal conductivity in the depths of the Earth. *Amer. J. Sci.*, **254**, 754–757.

Reed, G. W. & Turkevich, A., 1956. The uranium content of two iron meteorites. *Natl. Acad. Sci.—Natl. Res. Coun. Pub.* **400**, 97–99.

Ringwood, A. E., 1960. Some aspects of the thermal evolution of the Earth. *Geochim. et Cosmochim. Acta* (in press).

Russell, R. D., 1956. Interpretation of lead-isotope abundances. *Natl. Acad Sci.—Natl. Res. Coun. Pub.* **400**, 68–78.

Simon, F. E., 1953. The melting of iron at high pressures. *Nature* (*Lond.*), **172**, 746–747.

Strong, H. M., 1959. The experimental fusion curve of iron to 96 000 atmospheres. *J. Geophys. Res.*, **64**, 653–659.

Uffen, R. J., 1952a. A study of the thermal state of the Earth's interior. Ph.D. Thesis, Univ. Western Ontario, London, Canada).

Uffen, R. J., 1952b. A method of estimating the melting-point gradient in the Earth's mantle. *Trans. Amer. Geophys. Un.*, **33**, 893–896.

Verhoogen, J., 1956. *Physics and Chemistry of the Earth*, Vol. I, Chapter 2. (Pergamon Press, London.)

Wilson, J. T., 1957. Origin of the Earth's crust. *Nature* (*Lond.*), **179**, 228–230.

Wilson, J. T., 1959. Geophysics and Continental growth. *Amer. Sci.*, **47**, 1–24.

Wilson, J. T., 1960. Some consequences of expansion of the Earth. *Nature*, **185**, 880–882.

Zharkov, V. N., 1959. The fusion temperature of the Earth's mantle and the fusion temperature of iron under high pressures. *Bull* (*Izv.*) *Acad. Sci. U.S.S.R. Geophys. Ser.* 465–470 (English ed., 1960, pp. 315–320).

Heat Balance of the Earth's Core

John Verhoogen

"Das also war des Pudels Kern!"—*Goethe*, Faust, *Pt.I*, 1323.

Summary

It is shown that the heat necessary to maintain convection in the core may come from slow cooling and crystallization of the core with corresponding growth of the solid inner core. The present rate of cooling would be of the order of 10–45 degrees per 10^9 years, depending on temperature.

1. Introduction

Current views on the origin of the Earth's magnetic field require a convecting core; irregular variations of the length of the day provide evidence to the same effect. The necessary heat was originally thought to be radiogenic, but recent determinations of the radioactivity of iron meteorites (see, for instance, Reed 1959) suggest a rate of heat generation not greater than 1×10^{-18} cal/g sec, or $1{\cdot}9 \times 10^9$ cal/sec for the whole core, which would be about one hundred times too small (see below). Urey (1952) proposed gravitational energy as a source of heat, and Elsasser (1956) called upon chemical differentiation of an unspecified nature.

Bullard & Gellman (1954) have given likely values of the parameters relevant to the Earth's core, as well as estimates of the rate of dissipation of magnetic energy ($2{\cdot}1 \times 10^{10}$ cal/sec) and of the total heat flow at the core's surface ($3{\cdot}9 \times 10^{11}$ cal/sec). The assumed temperature is 5 000°C. If the inner and outer cores are assumed to have essentially the same composition (iron + nickel + silicon) the temperature at the inner core boundary (= melting point) is likely to be much less than 5 000°C (see Table 1). The purpose of this note is to recalculate the heat balance of the core on the basis of the melting-point data, and to show that slow cooling and progressive crystallization of the core could provide enough heat to maintain convection and the magnetic field.

2. The melting point of iron

Simon (1953) found that the change in melting point with pressure for many substances could be expressed by the relation

$$\frac{P}{a} = \left(\frac{T_m}{T_0}\right)^c - 1 \qquad (1)$$

where a and c are constants, T_m and T_0 are, respectively, the melting points at pressure P and at zero pressure. He suggested that the constants for iron should

be $a = 150\,000$ atm, $c = 4$. Gilvarry (1957) proposed $a = 347\,000$, $c = 1{\cdot}9$. More recently Strong (1959, 1960), having determined experimentally the m.p. up to about 100 000 atm, fitted his results to equation (1), obtaining $a = 75\,000$ atm, $c = 8$. Knopoff & MacDonald (1960) criticize Strong's value for c on the grounds that theoretically

$$c = \frac{6\gamma+1}{6\gamma-2}, \tag{2}$$

where γ is Grüneisen's ratio. For $c = 8$, $\gamma = 0{\cdot}405$, whereas the value of γ for iron at high pressure deduced from shock wave experiments is 1·6. This criticism is not fully justified, for equation (2) was derived (Salter 1954, Gilvarry 1957) in an attempt to explain Simon's empirical equation (1) on the basis of Lindemann's theory of melting. Lindemann's theory, as many others, suffers from the serious defect that it attempts to predict the m.p. entirely from properties of the solid phase, in spite of the fact that melting represents equilibrium between solid and liquid and must therefore depend also on properties of the liquid phase. Salter and Strong have both pointed out that experimental values of c do not in general agree with values derived from (2) in cases where γ may be determined directly from the expression

$$\gamma = \frac{\alpha V}{\beta_T c_v} = \frac{\alpha V}{\beta_s c_p} \tag{3}$$

where α is thermal expansion, V is volume, β_T and β_s are, respectively, the isothermal and adiabatic compressibilities, and c_v and c_p are, respectively, the specific heat at constant volume and pressure. In the following, calculations will be carried out for both Simon's and Strong's values of a and c. Significant values are listed in Table 1, in which the values of density and pressure are taken from Bullard's (1957) Model IV. This model was selected because, as will be shown later, the slope of the m.p. of iron at high pressure is such that the volume change on melting must be negligible; any model of the core which assumes a notable jump in density at the inner core boundary would be inconsistent with the present assumptions.

The melting temperatures listed in Table 1 represent maximum values as applied to the core, which presumably does not consist of pure iron. Additions of Ni and Si would probably depress the m.p. of iron at high pressure very much as they do at ordinary pressure, the depression being of the order of 300° for 28 per cent Si. The melting range (i.e. difference between liquidus and solidus) is small, as is the difference in composition of the solid and liquid phases in equilibrium.

3. Temperature gradient in the outer core

In the model here considered, the temperature gradient in the outer core must be greater than the adiabatic gradient but less than the melting-point gradient. The adiabatic gradient β is

$$\beta = \frac{\alpha}{c_p} gT = \frac{\gamma}{\phi} gT \tag{4}$$

where $\phi = v_p^2$, v_p being the velocity of the longitudinal elastic wave. The value of γ for liquid iron under conditions prevailing in the core is difficult to estimate, but as we are concerned with the maximum heat flow that need be maintained,

Table 1

Melting parameters for iron in the core, according to Simon and to Strong

Distance from centre (cm)	Density (g/cm³)	Pressure (bars)	Melting-point T_m (deg K)		dT_m/dP (deg/bar)		Melting-point gradient $-dT_m/dr$ (deg/km)	
			Strong	Simon	Strong	Simon	Strong	Simon
$3{\cdot}47 \times 10^8$	9·7	$1{\cdot}36 \times 10^6$	2610	3200	$2{\cdot}25 \times 10^{-4}$	$5{\cdot}38 \times 10^{-4}$	0·22	0·51
$1{\cdot}2 \times 10^8$	11·8	$3{\cdot}15 \times 10^6$	2890	3900	$1{\cdot}10 \times 10^{-4}$	$2{\cdot}78 \times 10^{-4}$	0·05	0·13
0	12·1	$3{\cdot}45 \times 10^6$	2940	3980	$1{\cdot}03 \times 10^{-4}$	$2{\cdot}76 \times 10^{-4}$	0	0

consideration of the melting point gradients, as given in Table 1, will be sufficient. At the outer boundary of the core, $\phi = 0{\cdot}66 \times 10^{12}\,\mathrm{cm^2\,s^{-2}}$, and $g = 10^3\,\mathrm{cm\,s^{-2}}$. The maximum gradient and other parameters concerned are then, for the two hypotheses considered:

Strong: gradient $< 0{\cdot}22$ deg/km, $\gamma < 0{\cdot}55$, $\alpha < 6{\cdot}2 \times 10^{-6}$/deg
Simon: gradient $< 0{\cdot}51$ deg/km, $\gamma < 1{\cdot}06$, $\alpha < 1{\cdot}2 \times 10^{-5}$/deg.

Bullard & Gellman (1954) evaluate the thermal conductivity k to be about 0·1 cal/cm deg s. The lower temperatures implicit in the present model reduces the electronic heat transfer, if the electrical conductivity is assumed to have the same value as in Bullard and Gellman's model; the phonon conductivity would be somewhat higher, but is quite uncertain, and the same value of k will be retained here. The heat flow q and the rate of heat loss for the entire core dQ/dt are then:

Strong: $q = 2 \times 10^{-7}$ cal/cm²s, $dQ/dt = 3 \times 10^{11}$ cal/s
Simon: $q = 5 \times 10^{-7}$ cal/cm²s, $dQ/dt = 7{\cdot}5 \times 10^{11}$ cal/s.

It is interesting to consider the conductivity required to carry this amount of heat, in a steady state, through the mantle. If the temperature at the surface $r = R$ is taken as 0, while T_c is the temperature at the core boundary, the gradient at $r = R_c$ is

$$\left(\frac{dT}{dr}\right)_{r=R_c} = -\frac{R}{R_c}\left(\frac{T_c}{R-R_c}\right)$$

which amounts to 1·45 deg/km for $T_c = 2\,300$°C (Strong) and 1·9 deg/km for $T_c = 3\,000$°C (Simon). The corresponding conductivities are 0·014 cal/cm sec deg (Strong) and 0·026 (Simon). The latter figure seems high.

4. Crystallization of the core

If the average temperature of the whole core decreases by an amount ΔT, crystallization will occur at the inner core boundary which will move outwards by a distance Δr such that

$$\frac{\Delta T}{\Delta r} = \frac{dT_m}{dr} = -\frac{dT_m}{dP}g\rho = \lambda,$$

say, and $\Delta r = \Delta T/\lambda$. The total amount of heat released is then

$$\Delta Q = [c_p M_c + 4\pi r_c^2 \rho \Delta H_m/\lambda]\Delta T \qquad (5)$$

where M_c is the mass of the whole core, r_c is the radius of the inner core, and ΔH_m is the heat of fusion (cal/g).

To evaluate the latter, we note that $\Delta H_m = T_m \Delta S_m$, ΔS_m being the entropy of melting. The entropy of melting for nearly all elements is close to 2–3 cal/g atom deg. Part of it arises from the volume expansion that occurs at the melting point:

$$\left(\frac{\partial S}{\partial V}\right)_T = \alpha/\beta_T.$$

For solids, such as iron, that melt with a small volume increase, this contribution to the entropy is small (about 0·24 cal/g atom deg for iron). Clearly, the greater

Table 1

Melting parameters for iron in the core, according to Simon and to Strong

Distance from centre (cm)	Density (g/cm^3)	Pressure (bars)	Melting-point T_m (deg K)		dT_m/dP (deg/bar)		Melting-point gradient $-dT_m/dr$ (deg/km)	
			Strong	Simon	Strong	Simon	Strong	Simon
$3{\cdot}47 \times 10^8$	9·7	$1{\cdot}36 \times 10^6$	2610	3200	$2{\cdot}25 \times 10^{-4}$	$5{\cdot}38 \times 10^{-4}$	0·22	0·51
$1{\cdot}2 \times 10^8$	11·8	$3{\cdot}15 \times 10^6$	2890	3900	$1{\cdot}10 \times 10^{-4}$	$2{\cdot}78 \times 10^{-4}$	0·05	0·13
0	12·1	$3{\cdot}45 \times 10^6$	2940	3980	$1{\cdot}03 \times 10^{-4}$	$2{\cdot}76 \times 10^{-4}$	0	0

consideration of the melting point gradients, as given in Table 1, will be sufficient. At the outer boundary of the core, $\phi = 0{\cdot}66 \times 10^{12}\,\mathrm{cm^2\,s^{-2}}$, and $g = 10^3\,\mathrm{cm\,s^{-2}}$. The maximum gradient and other parameters concerned are then, for the two hypotheses considered:

Strong: gradient $< 0{\cdot}22$ deg/km, $\gamma < 0{\cdot}55$, $\alpha < 6{\cdot}2 \times 10^{-6}$/deg
Simon: gradient $< 0{\cdot}51$ deg/km, $\gamma < 1{\cdot}06$, $\alpha < 1{\cdot}2 \times 10^{-5}$/deg.

Bullard & Gellman (1954) evaluate the thermal conductivity h to be about 0·1 cal/cm deg s. The lower temperatures implicit in the present model reduces the electronic heat transfer, if the electrical conductivity is assumed to have the same value as in Bullard and Gellman's model; the phonon conductivity would be somewhat higher, but is quite uncertain, and the same value of h will be retained here. The heat flow q and the rate of heat loss for the entire core dQ/dt are then:

Strong: $q = 2 \times 10^{-7}$ cal/cm²s, $dQ/dt = 3 \times 10^{11}$ cal/s
Simon: $q = 5 \times 10^{-7}$ cal/cm²s, $dQ/dt = 7{\cdot}5 \times 10^{11}$ cal/s.

It is interesting to consider the conductivity required to carry this amount of heat, in a steady state, through the mantle. If the temperature at the surface $r = R$ is taken as 0, while T_c is the temperature at the core boundary, the gradient at $r = R_c$ is

$$\left(\frac{dT}{dr}\right)_{r=R_c} = -\frac{R}{R_c}\left(\frac{T_c}{R-R_c}\right)$$

which amounts to 1·45 deg/km for $T_c = 2\,300$°C (Strong) and 1·9 deg/km for $T_c = 3\,000$°C (Simon). The corresponding conductivities are 0·014 cal/cm sec deg (Strong) and 0·026 (Simon). The latter figure seems high.

4. Crystallization of the core

If the average temperature of the whole core decreases by an amount ΔT, crystallization will occur at the inner core boundary which will move outwards by a distance Δr such that

$$\frac{\Delta T}{\Delta r} = \frac{dT_m}{dr} = -\frac{dT_m}{dP}g\rho = \lambda,$$

say, and $\Delta r = \Delta T/\lambda$. The total amount of heat released is then

$$\Delta Q = [c_p M_c + 4\pi r_c^2 \rho \Delta H_m/\lambda]\Delta T \qquad (5)$$

where M_c is the mass of the whole core, r_c is the radius of the inner core, and ΔH_m is the heat of fusion (cal/g).

To evaluate the latter, we note that $\Delta H_m = T_m \Delta S_m$, ΔS_m being the entropy of melting. The entropy of melting for nearly all elements is close to 2–3 cal/g atom deg. Part of it arises from the volume expansion that occurs at the melting point:

$$\left(\frac{\partial S}{\partial V}\right)_T = \alpha/\beta_T.$$

For solids, such as iron, that melt with a small volume increase, this contribution to the entropy is small (about 0·24 cal/g atom deg for iron). Clearly, the greater

part of the entropy of melting arises from structural changes. For instance, if each of the N particles of a solid of total volume V is considered to be effectively confined to a volume V/N in the solid state, but that all N particles share communally a volume V in the liquid state, positional degeneracy gives rise to a "communal" entropy $R = 1{\cdot}986$ cal/mole deg. Eyring's (Carlson, Eyring & Ree 1960) significant-structure theory of liquids proposes that actually only a fraction $N(V-V_s)/V$ of the particles share only a volume $(V-V_s)$ (where V is the volume of the liquid and V_s that of the solid at the m.p.); but the additional degrees of freedom (translational, rotational and vibrational) of these gas-like particles contribute a significant proportion of the total entropy of melting. As long as the solid and liquid states remain distinct, that is, barring a solid–liquid critical point, pressure should not have a large effect on such entropy terms, and it seems safe to predict that although the entropy of melting may decrease slowly with increasing pressure, it is unlikely to drop below $\frac{1}{2}R$. For iron this means that the heat of fusion ΔH_m is not likely to decrease; more probably, it will increase nearly proportionally to T_m. The heat of fusion of iron, which is 65 cal/g at ordinary pressure, would then be about 100 cal/g at the inner core boundary on Strong's data, and 135 cal/g if Simon's coefficients are used.

The mass of the core, in this model, is $1{\cdot}87 \times 10^{27}$ g, and c_p is taken as 0·18 cal/g deg. The first term on the right of (5) is $3{\cdot}37 \times 10^{26}$ cal/deg, while the second term is $4{\cdot}1 \times 10^{26}$ cal/deg (Strong) or $2{\cdot}2 \times 10^{26}$ (Simon). The total is thus $7{\cdot}5 \times 10^{26}$ or $5{\cdot}6 \times 10^{26}$. The present rate of cooling can then be calculated from the corresponding values of dQ/dt: it amounts to either 12 deg/10^9 years (Strong) or 45 deg/10^9 years (Simon). Equation (5) cannot be integrated directly to yield the age of the core, as λ and ρ depend on r, but a rough estimate of this age is about 3×10^9 years in both cases. The relative contribution of the latent heat of crystallization increases as the inner core grows, and its outward rate of growth decreases with time, if the heat flow remains constant.

Gravitational energy has not been considered, because it follows from the slope of the m.p. curve that the change in volume upon melting is insignificant at high pressure (less than $2{\cdot}3 \times 10^{-2}$ cm^3/mole, or 0·06 g/cm^3 under conditions prevailing in the inner core). Differences in composition between solid and liquid phases in the systems Fe–Ni and Fe–Si are small, and corresponding differences in density also appear negligible. Crystallization would thus not entail significant changes in gravitational energy. It would be interesting to know how large the density jump at the inner core boundary actually is, for it could provide a critical test of the liquid iron–solid iron hypothesis which would be ruled out if the density difference is of the order of 3 g/cm^3, or even 0·3 g/cm^3, as Bullen (1960) thinks it may well be.

The general conclusion then seems to be that, barring the occurrence of a notable density jump between outer and inner core, all data regarding the core, including melting point data on Fe and Ni, seem reasonably consistent. There is no difficulty in accounting for the heat sources necessary to maintain convection in the core, nor is there any major discrepancy that would require drastic revision of present views regarding its constitution.

University of California,
Berkeley,
California, U.S.A.:
1960, *August.*

References

Bullard, E. C., 1957. Gedenkboek F. A. Vening Meinesz (*Verh. Kon. Nederl.. Geol-Mijnbkdg. Gen.* **18**), 23–41.

Bullard, E. C. & Gellman, H., 1954. *Phil. Trans. Roy. Soc. London*, **A 247**, 213–278.

Bullen, K. E., 1960. In press and personal communication.

Carlson, C. M., Eyring, H. & Taikyue Ree, 1960. *Proc. Nat. Acad. Sci.*, **46**, 649.

Elsasser, W. M., 1956. *Rev. Mod. Phys.*, **28**, 135–163.

Gilvarry, J., 1957. *J. Atm. Terr. Phys.*, **10**, 84–95.

Knopoff, L. & Macdonald, G. J. F., 1960. *Geophys. J.*, **3**, 68–77.

Reed, G. W., 1959, in *Researches in Geochemistry* (edit. by P. H. Abelson) (Wiley, New York).

Salter, L., 1954. *Phil. Mag.*, **45**, 369–378.

Simon, F., 1953. *Nature* (*Lond.*), **172**, 746.

Strong, H. M., 1959. *J. Geophys. Res.*, **64**, 653–660.

Strong, H. M., 1960. *Amer. Scientist*, **48**, 58–79.

Urey, H. C., 1952. *The Planets* (Yale Univ. Press, New Haven).

The Flow of Heat through the Floor of the Atlantic Ocean

E. C. Bullard and A. Day

"In his master's steps he trod . . .
Heat was in the very sod"—*Traditional*

Summary

Fifteen measurements of the heat flow through the ocean floor have been made in the Atlantic and one in the Mediterranean. One measurement in the central valley of the Mid-Atlantic Ridge gives a heat flow of 6·5 μcal/cm^{2}s. Similar very high values have been found by others on the crest of the East Pacific Rise. The possibility that these high values represent the rising limb of a convection current is discussed. The mean of the remaining fifteen values and of five from a previous investigation, give a mean heat flow of $1{\cdot}06 \pm 0{\cdot}055$ μcal/cm^{2}s which is very close to the mean for all measurements in the Pacific excluding those on the East Pacific Rise.

The equipment has been modified so that the temperature gradients in the whole probe and in its lower half can be determined separately. The lower half gives a value 8·8 per cent above that for the whole probe, the reason for this is not known.

The thermal resistivity of ocean sediment, R, is found to be a linear function of the water content, w (in per cent water of the wet weight). At the temperature of the ocean floor the relation is

$$R(\text{cms deg C/cal}) = (168 \pm 14) + (6{\cdot}78 \pm 0{\cdot}31)w.$$

1. Introduction

Five measurements of the flow of heat through the floor of the Atlantic Ocean obtained in 1952 have been described in a previous paper (Bullard 1954). This work has been continued by the same methods during voyages in *Discovery II* in 1954, 1956 and 1958. Sixteen further measurements have been made. The purpose of this paper is to describe this work.

2. Measurement of the temperature gradient

As before the temperature gradient in the sediment beneath the ocean floor was measured by a probe 4·6 m long and 2·7 cm in diameter. The probe used in 1954 had two sets of junctions, each set having three pairs of junctions in series. One set measured the temperature difference between points near the top and bottom of the probe and the other that between the top and a point about half-way

down. In 1956 and 1958 the second set of junctions measured the temperature difference between the bottom of the probe and a point 276 cm higher up. The latter arrangement is much to be preferred since a set of junctions in the lower half of the probe will give a satisfactory measurement even if the probe does not penetrate completely. If penetration is complete a check is obtained both on the linearity of the relation between temperature and depth and on the functioning of the apparatus. The two sets of junctions will be called the *long* and the *short* junctions.

An attempt was made to reduce the time needed to attain equilibrium by putting the junctions at the lower end of the probe in a tube 4 mm in diameter extending 13 cm beyond the end of the 2·7 cm tube and the top junctions in a similar tube running parallel to the upper part of the probe and distant 10 cm from it. This arrangement was not satisfactory since one or both of the narrow tubes was always broken, probably when pulling the probe out of the sediment. Also the temperature rise due to friction while the tube was penetrating the sediment was much greater than with a larger tube. In one lowering the initial rise obtained by extrapolation back to the instant of penetration as previously described (Bullard 1954) was 9·1 deg C compared to a mean of 0·8 deg C and a maximum of 1·6 deg C for the 2·7 cm probe. If the friction is proportional to the circumference of the probe and all the heat went into the probe, the temperature rise would be inversely proportional to the probe diameter; a large rise is therefore to be expected for the 4 mm tube. In view of the difficulties experienced, the use of this type of probe was abandoned. All the results given in this paper have been obtained with the probe 2·7 cm in diameter.

Wind tunnel tests at the National Physical Laboratory on a twice full size model of the probe showed that in free fall in water it is stable for inclinations up to 10° from the vertical; at inclinations beyond this it becomes unstable and sets itself broad-side on to the direction of motion. In practice the probe is attached to a cable at its upper end which will increase its stability. Although there seems little risk of it overturning, fins have been added above the recording case which render it completely stable. A device that released the probe when its lower end was from 0 to 4 m above the sea floor and allowed it to fall freely was used at five of the stations in 1956.

The recorder was similar to that previously described except that an additional relay switched the recording galvanometer from the long to the short set of junctions and back again about once every two minutes. An inclinometer, recording on the same film as the galvanometer, recorded the inclination about one horizontal axis as a function of time. This showed that the difference between the attitude of the probe when hanging in the water and when in the sediment averaged 4°, the largest inclination observed was 9°. This has a negligible effect on the measurements and there is little to be gained by fitting a second inclinometer to record tilts about an axis perpendicular to the first.

3. Thermal conductivity

Mr D. W. Butler and Mr E. H. Ratcliffe of the National Physical Laboratory have made 44 determinations of the thermal conductivity and water content of ocean sediments; 15 of these values for globigerina ooze have previously been published (Bullard 1954, Table 4), the remaining 29 are given in Table 1. The specimens of red clay are from Revelle and Maxwell's (1952) stations, the first

3 specimens of mud are from Mission Bay, a tidal mud flat near San Diego, the next 2 from the San Diego trough and the remaining 8 specimens are of a black mud from our Mediterranean station 8 (Table 2). The specimens of impure globigerina ooze are from our stations 6, 7 and 8. The water contents vary from 20 to 69 per cent, the extreme values having been obtained by drying and adding water to the material from the cores.

Table 1

Thermal conductivity of sediments

Water content (% wet wt)	% $CaCO_3$ on dry wt	Density g/cm³	Thermal conductivity cal/cm s°C $\times 10^{-4}$
Globigerina ooze			
37·8	56	1·54	23·3
43·8	48	1·47	20·7
43·4	,,	1·47	22·2
36·9	59	1·55	25·5
38·5	60	1·54	25·2
Red Clay			
52	0·5	1·43	19·3
54	,,	1·39	19·3
56·5	,,	1·38	19·3
50	2·8	1·47	21·7
50	,,	1·47	22·0
42·5	0·5	1·58	23·7
43·5	,,	1·57	24·3
52·5	0·7	1·41	19·1
52	0·3	1·40	19·6
69·5	0·7	1·20	16·8
61·8	0·6	1·27	17·3
Dark grey mud			
55	10·1	1·32	19·1
52·5	,,	1·36	19·0
56·5	,,	1·31	18·8
51·5	11·2	1·37	19·4
46	,,	1·47	20·6
46·8	29	1·43	20·8
45·7	,,	1·45	21·7
44·4	20	1·47	22·4
44·6	25	1·47	22·4
42·6	21	1·49	22·4
38·0	24	1·57	23·9
37·8	,,	1·57	23·0
38·9	23	1·56	24·4

The ditto marks in column 2 indicate that a single determination of the $CaCO_3$ content is believed to represent two or more specimens used for the determination of conductivity.

A theoretical expression for the variation of conductivity with water content has been compared with these observations (Bullard, Maxwell and Revelle, 1956) and a straight line fitted over the range of practical interest. It has been found,

however, that the reciprocal of the conductivity (the resistivity) gives a straight line over a larger range of water content (Figure 1). A least squares fit gives

$$R = (161 \pm 14) + (6{\cdot}51 \pm 0{\cdot}30)w$$

where w is the water content in per cent of water on the wet weight and R is the resistivity in deg C cm s/cal. The standard deviation of one measurement is 17 deg C cm s/cal or 4 per cent. The residuals show no significant curvature in the relation between w and R.

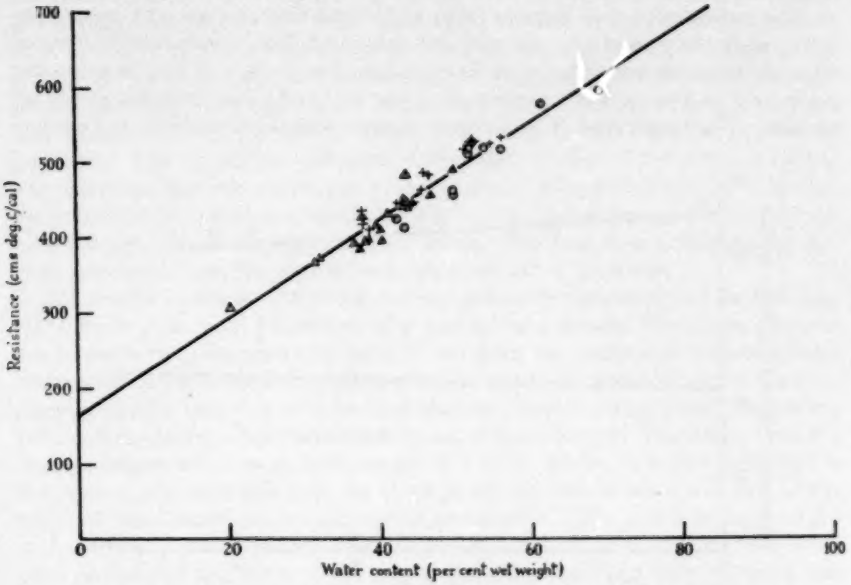


FIG. 1.—Relation between thermal resistivity (laboratory conditions) and water content (*per cent* of water on the wet weight). ⊙ red clay, △ globigerina ooze, + mud.

An examination of the residuals from (1) shows that the resistivity of red clay and globigerina ooze of a given water content do not differ significantly; that of mud is 2 per cent higher, but the difference is barely significant and has been ignored.

The conductivity at the temperature and pressure of the ocean floor is not the same as that measured in the laboratory at a temperature of about 25 deg C and a pressure of one atmosphere. The temperature coefficient of the thermal conductivity of ocean sediments has been determined by Butler (Von Herzen & Maxwell 1959, Ratcliffe 1960); he finds that the conductivity decreases by 6 per cent between 25 and 4 deg C. Pressure is believed to increase the conductivity by about 1 per cent for every 1 000 fm of depth. At the depths of our stations a uniform correction of 4 per cent is adequate and the resistivity has been taken as

$$R = 168 + 6{\cdot}78w \qquad (1)$$

The resistivities for the five stations measured in 1952 have been recalculated from

this relation. The results for these stations given in Figure 3 therefore differ slightly from those given previously (Bullard 1954).

The conductivity required in a heat flow measurement is not the arithmetic mean over the length of the temperature probe, but the reciprocal of the mean resistance; since the relation (1) is linear the mean resistance can be found directly from the mean water content of a core. This has been found by averaging the water contents obtained on 6-inch lengths of core.

Unfortunately in 1954 and 1956 we were unable to obtain cores covering the full 4·6 m length of the probe. It is necessary therefore to consider the error incurred if the conductivity is deduced from a core that is shorter than the probe.

The individual 6-inch samples taken at the heat flow stations have water contents varying from 23 to 64 per cent, but this variability is considerably reduced when the water content is averaged down a core. Ten cores with lengths exceeding 2·4 m gave an average water content of 43 per cent and all lay between 40 and 48 per cent. The variation of the mean water content with depth is shown in Figure 2.

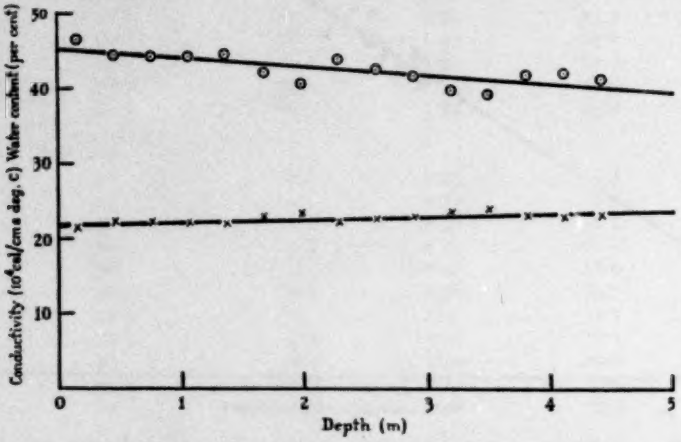


FIG. 2.—Variation of water content ⊙ and conductivity × with depth, mean of 10 cores.

Here all 6-inch samples from the 10 cores lying in 1 ft depth ranges have been averaged. The corresponding conductivities are also shown.

If the conductivity for the long set of junctions is estimated from the mean water content between depths of 0·3 and 2·4 m, it will on the average be underestimated by 1·3 per cent; that of the short set in the bottom half of the probe will be underestimated by 2·2 per cent. There will also be a random error in estimating the conductivity of the missing lower half of a core; this will be less than the standard deviation of the conductivity of the 10 cores from this mean which is 4 per cent. No correction has been made for these errors which are negligible. The smallness of the variability in the mean water contents of the cores is probably due to the uniformity of the sediments. The probe would not penetrate the sandy sediments of the abyssal plains and our stations are all in situations where softer sediments could be found; in the Atlantic they are mixtures of globigerina ooze and clay and at the single station in the Mediterranean a soft black mud. The above average water contents could not be expected to apply to other types of sediment.

4. The measurements

In 1954 the temperatures were recorded for 22 minutes, this gave an undesirably large correction for lack of equilibrium (mean 29 per cent). The interval during which measurements were taken was therefore increased to between 40 and 50 minutes in 1956 and 1958, this reduced the mean corrections for these years to 19 and 15 per cent. The correction was determined by fitting the theoretical relation between temperature and time (equation 7 of Bullard 1954) to the observations by least squares. With one exception the fit was satisfactory, the standard deviation of a measured temperature being 0·0044 deg C. This is several times larger than in the measurements made in 1952, the increase is due primarily to a deterioration in the galvanometer mirrors. If the measurements were independent the standard error of the extrapolated equilibrium gradient would be about 2 per cent, but this is probably an underestimate since the errors at successive measured points on a record tend to be correlated (as can be seen in Figure 3 of Bullard & others 1956). The record at station 17 did not give a satisfactory record from the long set of junctions. The uppermost junctions were suddenly heated by about 0·1 deg C after the probe had been 16 minutes in the sediment; presumably the ship pulled on the probe and bent it somewhere below the top junction causing frictional heating near the top, but no disturbance lower down. The heat flow at this station has been calculated from the results from the short set of junctions.

On several occasions the probe did not penetrate completely. There is then some doubt as to what proportion of it was in the sediment, though an estimate can be made from the mud sticking to it and from the position of the sharp bend which is usually produced on pulling out. In 1954 there was only one such station (number 6); the heat flow at it has been deduced from the long junction assuming 158 cm of the probe to have been sticking out of the sediment. The results from the short junctions, which were in the upper half of the probe, have not been used at this station. In 1956 and 1958 the short junctions were in the lower part of the tube and were unaffected by incomplete penetration. The heat flow adopted for such stations is that obtained from the short junctions. For stations with complete penetration the mean of the results from the short and long junctions was adopted. The results for the long and short sets of junctions at 9 stations with complete penetration show a systematic difference. At every station the temperature gradient is less for the long set of junctions than for the short set. The difference varies from 0·3 to 18 per cent and averages 8·8 per cent. A small difference in the opposite direction might have been expected in the measurements made in 1956 and 1958, since the water content, on the average, decreases with depth and the conductivity increases; this should cause the short set of junctions to give a gradient 0·9 per cent smaller than the long set. In 1954, when the short junctions were in the upper half, the observed differences were in the same direction as 1956 and 1958, but were only 0·3 and 2·4 per cent and not significant. The observed differences are also in the opposite direction from those that would be produced by chemical or bacterial heating near the surface. The cause of the discrepancy has not been established, the most likely explanation seems to be that there is some movement of water in the top metre or two of sediment and that this transports a small part of the heat and reduces the gradient near the surface.

There are 5 stations with incomplete penetration at which results were obtained from both sets of junctions. Four of these give a larger gradient over the whole probe than over the bottom half, this is the reverse of the behaviour found at the stations with complete penetration. One of these stations (number 12 in Table 2)

Table 2

Measurements of heat flow

Station number	Latitude N	Longitude W	Depth m	Long Junction: Temp. gradient °C/km	Long Junction: Thermal cond. cal/cm s °C ×10⁻⁴	Long Junction: Heat flow μcal/cm² s	Short junctions: Temp. gradient °C/km	Short junctions: Thermal cond. cal/cm s °C ×10⁻⁴	Short junctions: Heat flow μcal/cm² s	Adopted heat flow μ cal/cm² s
				1954 November						
6	39° 36′	12° 13′	3020	45·9*	23·0	1·06*	—	—	—	1·06
7	35 59	9 59	4534	37·4	23·1	0·87	37·5	23·1	0·87	0·87
8	35 58	4 34	1251	56·6	21·3	1·21	58·0	21·3	1·24	1·22
				1956 July–August						
9	45 28	5 47	4592	31·5	22·6	0·71	34·5	22·6	0·78	0·75
10	46 32	13 04	4413	45·7	21·6	0·98	55·0	21·8	1·20	1·09
11	46 30	22 58	4084	53·5	22·2	1·19	61·4	22·8	1·40	1·29
12	46 37	27 18	4109	415·8*	20·7	8·61*	314·9	20·7	6·52	6·52
13	36 20	21 00	4844	51·2	21·2	1·08	56·7	21·2	1·20	1·14
14	35 36	19 02	5375	68·3*	20·5	1·40*	66·6	20·1	1·34	1·34
15	35 34	18 56	5380	44·0*	20·5	0·90*	36·4	20·1	0·93	0·93
16	36 39	17 21	5146	51·0	21·3	1·09	55·5	21·3	1·18	1·14
17	44 55	10 45	4844	†	21·5	†	63·6	21·8	1·39	1·39
				1958 June–July						
18	40 59	15 09	5305	52·8*	22·8	1·20*	49·0	23·2	1·14	1·14
19	42 18	11 53	3063	33·0	21·7	0·72	38·8	21·9	0·85	0·78
20	41 27	14 40	5260	61·6*	21·6	1·32*	55·2	21·8	1·21	1·21
21	43 42	12 39	5030	50·5	22·8	1·15	50·7	23·0	1·17	1·16

* Incomplete penetration.
†Disturbed.

is exceptional in having a very high heat flow and a large difference between the two gradients. The mean discrepancy for the other 4 is 4·1 per cent, this is probably due to an underestimate of the penetration, it is quite possible that the sharp bend in the probe on pulling out usually occurs somewhat below the surface and also that the topmost unconsolidated sediment is washed off the probe on the way up to the sea surface. An underestimate of penetration by 30 cm would produce overestimate of the gradient by 7 per cent.

5. Discussion

The results are given in Table 2 and Figure 3. The figure also shows the revised results at the five stations from Bullard (1954) and therefore contains all

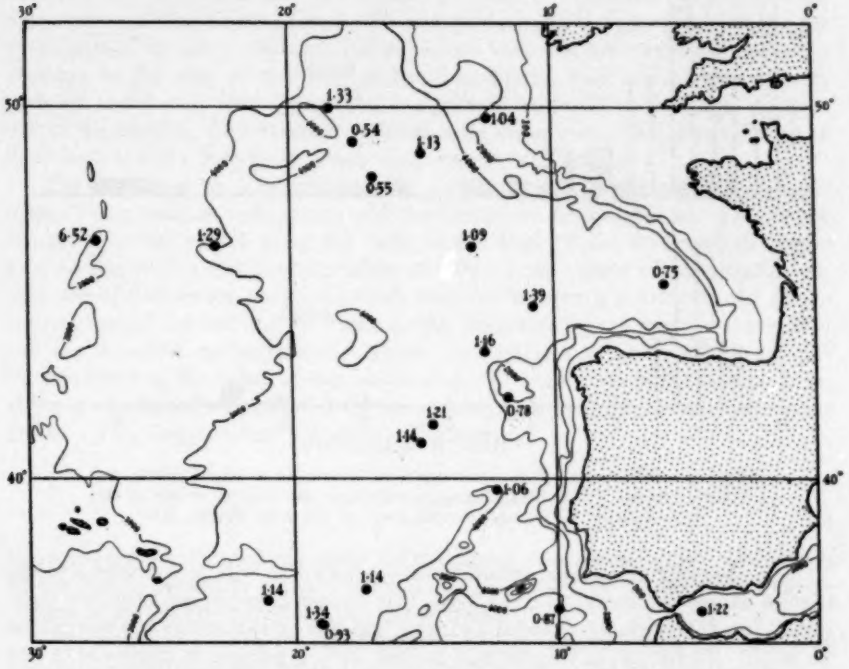


FIG. 3.—Heat flow in μcal/cm^2 s, contours in metres.

the measurements made from *Discovery II*. Preliminary results for stations 6, 7 and 8 were given by Bullard, Maxwell & Revelle (1956), these are superseded by those of Table 2. All heat flows in this paper are in μcal/cm^{2}s, this unit has not been repeated after each value referred to in the text.

The value of 6·5 μcal/cm^{2}s found at station 12 is 4·7 times the largest value found at any other station and is clearly exceptional. This value was found from the short set of junctions. The long set is believed not to have penetrated completely and to have projected 84 cm above the sediment; if this is correct the heat flow derived from it is 8·6, whilst if it did penetrate completely it would be

7·0. The reality of the high heat flow is therefore confirmed by the long set of junctions.

The station is in the central valley of the Mid-Atlantic Ridge (Heezen & others 1959, Hill 1960). It is most desirable that further measurements should be made both in the valley and on lines crossing the ridge. This will not be easy since the area is rocky and it is a matter of chance whether the probe hits a pocket of sediment or a rock. Two galvanometers were broken and a coring tube bent in efforts to repeat the measurement at station 12. Similar high heat flows (up to 8·1) have been found by expeditions from the Scripps Institution on the crest of the East Pacific Rise (Bullard, Maxwell & Revelle 1956, von Herzen 1959). A histogram of these values is shown in the upper section of Figure 4. It seems probable that very high values of the heat flow are characteristic of the crests of mid-ocean ridges.

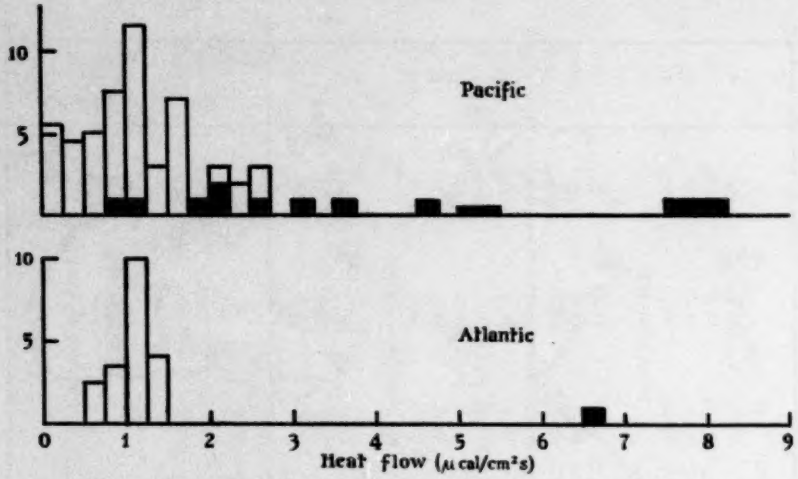


FIG. 4.—Histogram of measurements of heat flow in the Atlantic and Pacific Oceans. The black bars are values in the central valley of the Mid-Atlantic Ridge and on the crest of the East Pacific Rise.

The observed temperature gradient of 315 deg C/km will only continue to the bottom of the sediments. In the lavas that almost certainly underlie the sediments, it will be reduced to about 140 deg C/km if the conductivity is 0·005 cal/cm sec deg C. These values suggest that molten rock is present at depths of a few kilometres below the ocean floor. The origin of the Mid-Atlantic Ridge has been discussed by Hess (1954) and by Heezen, Ewing & Tharp (1959). It forms part of a worldwide system of connected ocean ridges, one branch of which runs into the Red Sea. Since the Red Sea has a central valley which is believed to be a tensional crack and to be floored by a dyke (Girdler 1958), it is natural to suggest that the central valleys of other parts of the system, and in particular the central valley in which our station 12 lies, are similar features. Such a world encircling tensional feature might be due to convection currents in the mantle rising under the ridges, spreading out sideways and sinking again. The horizontal limbs of the currents would produce the forces that open the tensional crack and the rising current would bring hot material near the surface and produce vulcanism and a high heat flow.

Such a scheme is speculative, but it is worth considering what other observable phenomena it might produce. A positive gravity anomaly might be expected over the valley due to the intrusion of dense rock into the opening crack (as in the Red Sea) but this should be superposed on a regional low associated with the rising current. The orders of magnitude of the quantities involved have been considered by Bullard & others (1956) who conclude that if the temperature in the rising current is one or two hundred degrees above the normal value an anomaly of -10 to -30 mgal would be expected. No gravity measurement appears to have been made over the central valley of any ridge. The values on the ridge are very scattered but do not show any consistent tendency to be below the values in the neighbouring areas.

If the rising current under the ridge turned down again on each side, a low heat flow would be expected over the descending limbs. The stations in Figure 3 do not show strips of low heat flow, but are so sparsely distributed that if they exist they could easily have been missed. There are values of 0·54 and 0·55 about 700 km to the east of the crest. Von Herzen has found values as low as 0·14 in the basin 1 600 km to the west of the crest of the East Pacific Rise and a value of 0·23 1 000 km to the east. He suggests that the pattern he has found is due to convection in the mantle. Clearly much remains to be done to elucidate the meaning of these high and low heat flows which were entirely unexpected.

The remaining 20 observations give a symmetrical histogram with a single peak. Their mean is $1{\cdot}06 \pm 0{\cdot}055$ and their standard deviation 0·26. The Pacific results, with the values along the crest of the East Pacific Rise excluded, give $1{\cdot}09 \pm 0{\cdot}094$ with a standard deviation of 0·65. If all values within 2 000 km on each side of the rise are excluded (which excludes the very low values in the neighbouring basins but not the low values in the Chile trench) the mean is $1{\cdot}17 \pm 0{\cdot}11$ and the standard deviation 0·60. There thus appears to be no significant difference between the mean of our small sample of heat flows in the eastern basin of the north Atlantic and that of the larger sample drawn from a wide area in the Pacific. This suggests that the mean of both sets

$$1{\cdot}08 \pm 0{\cdot}054\ \mu\text{cal/cm}^2\text{s}$$

has some real significance as a mean for the oceans away from ridges. The spread of the Pacific results is greater, probably because they are drawn from a wider variety of environment including an ocean trench.

The results away from the ridge do not show any obvious trends. It is possible that their variation from 0·54 to 1·39 μcal/cm^{2}s is controlled by local factors such as irregularities in the buried rock surface beneath the sediment and that the station spacing is too wide to reveal the real pattern. The large difference between stations 14 and 15 which are only 7 km apart and give heat flows of 1·39 and 0·93 is curious and unexplained.

Acknowledgments

We are indebted to the Director of the National Institute of Oceanography for allowing us to use *Discovery II* for these measurements and to the Captain, officers and crew of the ship and to the staff of the Institute for much assistance in the troublesome operation of the equipment. We have also had much help from the

staff and students of this Department, especially from Dr M. N. Hill, Mr M. Keen and Dr T. Allen. We are also indebted to Mr D. W. Butler and Mr E. H. Ratcliffe of the National Physical Laboratory for their measurements on the thermal conductivity of our specimens.

Department of Geodesy and Geophysics,
Cambridge:
1960 *May.*

References

Bullard, E. C., 1954. *Proc. Roy. Soc. A.*, **222**, 408–429.
Bullard, E. C., Maxwell, A. E. & Revelle, R., 1956. *Adv. Geophys.*, **3**, 153–181.
Girdler, R. W., 1958. *Quart. J. Geol. Soc. Lond.*, **114**, 79–105.
Heezen, B. C., Tharp, M. & Ewing, M., 1959. The Floors of the Oceans, I. The North Atlantic, *Geol. Soc. Amer. Sp. Pap. 65.*
Hess, H. H., 1954. *Proc. Roy. Soc. A.*, **222**, 341–348.
Hill, M. N., 1960. *Deep-Sea Res.*, **6**, 193–205.
Ratcliffe, E. H., 1960. *J. Geophys. Res.*, **65**, 1535–1541.
Von Herzen, R., 1959. *Nature* (*Lond.*), **183**, 882–883.
Von Herzen, R. & Maxwell, A. E., 1959. *J. Geophys. Res.*, **64**, 1557–1563.

VI. Composition

F. Birch, Composition of the Earth's mantle 295

A. L. Hales, A weak layer in the mantle? 312

F. F. Evison, Rock magnetism in Western Europe as an indication of continental growth 320

Composition of the Earth's Mantle

Francis Birch

"There is something fascinating about science. One gets such wholesale returns of conjecture out of such a trifling investment of fact."—*Mark Twain*, Life on the Mississippi.

Summary

Measurements of the velocity of compressional waves in silicates and oxides having a range of density from 2·6 to 5 g/cm³ suggest a simple dependence of velocity upon density and mean atomic weight. Consequences of the assumption that this relation remains valid at high pressures are examined with reference to the composition of the mantle, especially of the transition layer. It appears probable that the abnormally high rate of increase of velocity with depth in the transition layer may be accounted for principally in terms of phase change, with little change of composition. Recent studies with strong shock waves are examined in connection with the composition of the core; the evidence is unfavourable to the hypothesis of a metallic core of light elements, but is consistent with a core of iron alloy.

1. Introduction

Recent measurements of the velocity of compressional waves in rocks suggest a new approach to the problem of interpreting the seismic velocities of the Earth's mantle in terms of chemical composition and crystal structure. These measurements lead to a tentative generalization that, as a first approximation, the compressional velocity (V) in silicates and oxides is determined by two principal variables, density and mean atomic weight. A dependence upon density alone has sometimes been assumed for the correlation of seismic velocity and gravity, especially for near-surface layers. The validity of this is now seen to require the assumption that these layers consist mainly of rocks similar in composition to those common at the surface. As the composition of the deeper interior is still uncertain, and as the seismic velocities are the only properties of the interior which are fairly directly determined, it is essential to extract as much information as possible from observable correlations of velocity with other factors.

The details of the measurements of velocity, at pressures up to 10 kilobars, have been given elsewhere (Birch 1960). Above several kilobars, most of the disturbing effects of porosity are suppressed for the igneous and metamorphic rocks and certain regularities begin to appear. When the velocities at 10 kilobars are plotted against density, the effect of heavy-metal content is apparent: for a given velocity, iron-rich or titanium-rich samples have higher densities than do samples low in these elements. At a velocity of about 7 km/s, the observed range of

density is from 2·8 to 5 g/cm³. Most of the measurements are for iron-poor rocks, and with some selection, these have been used to determine a linear relationship between velocity and density for a constant mean atomic weight (m) of about 21. It is then possible to interpret the other more scattered points in terms of a series of lines of constant mean atomic weight parallel to but displaced from the well-determined line for $m \approx 21$.

The mean atomic weight of a compound is the molecular weight divided by the number of particles in the molecular formula; for example, the molecular weight of SiO_2 is 60·09, the number of particles 3, and m is 20·03. The composition of rocks is generally expressed in proportions by weight of the oxides; if x_i is the proportion by weight and m_i the mean atomic weight of oxide i, the mean atomic weight of the rock is $m = (\Sigma x_i/m_i)^{-1}$. In these calculations the hydroxyl group $(OH)^-$ is counted as a single particle. Values for a few common minerals are shown in Table 1, and for various collections of rock analyses in Table 2. The mean atomic weight is a significant parameter for a number of physical properties (Birch 1952, pp. 238, 256). In the context of geochemical abundances, it is evidently an index of heavy metal, especially of iron content, and in the following discussion of major divisions of the mantle the heavy metal content will be identified with iron content.

The velocity–density line for $m = 21$ is based on measurements for 47 materials, most of which were represented by three orthogonal specimens; these include granites, dunites, anorthosites, pyroxenites, jadeite and corundum, plus a number of metamorphic rocks of low iron content; values for periclase, spinel and corundum formed from single-crystal elastic constants have been included. Analyses are available for some of these materials; the mean atomic weight is, of course, not exactly the same for all of these but probably is always between 20 and 22, with 21 the best round number as an average. The equation of the least-squares line is $V = -2{\cdot}55 + 3{\cdot}31\rho$, where V is in km/s and ρ in g/cm³. The standard deviation of a single observation is 0·28 km/s. The correlation coefficient for velocity is 0·97; thus individual differences of crystal structure and composition have relatively little influence.

The density of any material depends upon two factors—the intrinsic mass of the particles, or mean atomic weight, and the closeness of packing, or mean atomic volume. In all the silicates and oxides, the oxygen ions account for most of the volume. Along a line of constant mean atomic weight, density can change only through changes of mean atomic volume, reflecting contractions of the crystalline lattice. In the Earth, such changes are produced mainly by pressure, first as ordinary compression without change of crystal structure, eventually as changes of crystal structure to more tightly packed lattices. Probably all the loosely-packed silicates undergo changes of structure at high pressures; known transformations include quartz to coesite, albite to jadeite plus quartz, anorthite plus portlandite to zoisite. The more tightly packed olivines are transformed to spinels at pressures of 50 kilobars or more (Ringwood 1958a; Dachille & Roy 1960). For the present purpose, the question of interest is the change of velocity which accompanies such transformations. The experimental data suggest that, regardless of crystal structure, the representative point on the velocity–density diagram will move along a line of constant mean atomic weight, essentially parallel to the line determined by the observed values for the iron-poor rocks. Thus albite and jadeite are related in this way; also, with allowance for scatter of compositions and measurements, the gabbros and eclogites. As a known exception, α-quartz has a higher density

Table 1

Mean atomic weight of minerals

Mineral	Ideal composition	*m*
Chrysoberyl	$BeAl_2O_4$	18·14
Spodumene	$LiAlSi_2O_6$	18·60
Brucite	$Mg(OH)_2$	19·44
Serpentine	$Mg_3Si_2O_5(OH)_4$	19·80
Chlorite	$Mg_5Al_2Si_3O_{10}(OH)_8$	19·85
Talc	$Mg_3Si_4O_{10}(OH)_2$	19·96
Quartz	SiO_2	20·03
Enstatite	$MgSiO_3$	20·08
Forsterite	Mg_2SiO_4	20·10
Pyrope	$Mg_3Al_2Si_3O_{12}$	20·15
Periclase	MgO	20·16
Albite	$NaAlSi_3O_8$	20·17
Jadeite	$NaAlSi_2O_6$	20·21
Kyanite	Al_2SiO_5	20·25
Spinel	$MgAl_2O_4$	20·32
Corundum	Al_2O_3	20·39
Muscovite	$KAl_3Si_3O_{10}(OH)_2$	20·96
Orthoclase	$KAlSi_3O_8$	21·40
Anorthite	$CaAl_2Si_2O_8$	21·40
Zoisite	$Ca_2Al_3Si_3O_{12}(OH)$	21·63
Diopside	$CaMgSi_2O_6$	21·65
Grossularite	$Ca_3Al_2Si_3O_{12}$	22·52
Almandite	$Fe_3Al_2Si_3O_{12}$	24·88
Andradite	$Ca_3Fe_2Si_3O_{12}$	25·40
Ferrosilite	$FeSiO_3$	26·38
Rutile	TiO_2	26·63
Fayalite	Fe_2SiO_4	29·11
Ilmenite	$FeTiO_3$	30·35
Hematite	Fe_2O_3	31·94
Magnetite	Fe_3O_4	33·08

Table 2

Mean atomic weight for certain groups of rocks

Designation	*m*	Reference
All sediments	20·8	Poldervaart 1955, p. 132
546 granites	20·9	Daly 1933, p. 9
Dunite, Twin Sisters Mt.	20·9	Ross & others, 1954, p. 707
Peridotite nodules (2)	21·0	Hess in Poldervaart 1955, p. 134
Average igneous rock	21·4	Clarke 1924, p. 29
9 dunites	21·5	Nockolds 1954, p. 1023
23 peridotites	21·6	Nockolds 1954, p. 1023
Silicate phase, 94 chondrites	21·8	Urey & Craig 1953, p. 55
Lithosphere	21·8	Poldervaart 1955, p. 133
637 mafic igneous rocks	22·0	Nockolds 1954, p. 1032
14 picrites	22·1	Daly 1933, p. 19
34 eclogites	22·1	Lapadu-Hargues 1953, p. 166
Hortonolite dunite, Bushveld	24·3	Hall 1932, p. 325

but lower velocity than β-quartz, though the changes are small. It will be supposed that normally the denser phase has the higher velocity, according to a relationship of the form $V = a(m)+3{\cdot}31\rho$. The function $a(m)$ is best determined for $m \approx 21$; a small group of points at $m \approx 24$ gives another value, and finally a scattered group of iron and titanium oxides indicate roughly the value for $m \approx 30$. Estimates of $a(m)$ are shown in Table 3; values for m greater than about 25 are not required for the discussion of the mantle. We shall wish, however, to use velocity to determine density, and for this purpose write $\rho = A(m)+0{\cdot}302\,V$, where $A(m) = -0{\cdot}302a(m)$. Again, several exceptions may be noted: rutile, grossularite, and calcium-rich minerals in general have somewhat higher velocities than are found from their densities with these relations. These complications will be ignored for the present. It will also be supposed that changes of temperature leave the velocity–density relation unchanged; the error so introduced will be examined below.

As a provisional hypothesis, it will be postulated that the velocity–density relations, as expressed above, hold for all changes of density in the Earth's mantle, however produced. Subject to the limitations of the data, this provides a new test for homogeneity: for a homogeneous, self-compressed layer, the velocity–density curve will be a straight line parallel to the lines of constant m. This condition is necessary but evidently not sufficient for true chemical uniformity, since variations of the proportions of the light elements have little effect on m and hence on the relations between the physical properties. So far as velocity–density relations are concerned, "homogeneous" means only that the iron content is uniform. Rocks as different chemically as granites and dunites have nearly the same mean atomic weight; the implication is that at the same density they have the same velocity. We conclude at once that little can be discovered about the distribution of the light elements from the seismic velocities. At best, some indication of iron content may be obtainable.

2. Density distributions for the mantle

Density distributions have been given by Bullen, Bullard, Ramsey, Bolt and others, each corresponding to a set of assumptions, some physical, some purely mathematical. Most of these bear some relation to a particular set of velocities. Thus the velocity distribution becomes a depth scale on the velocity–density plot. A curve for each density distribution may be plotted on the velocity–density diagram, and the position of the derived curve in relation to the experimental points may be taken as an indication of the mean atomic weight or iron content implied in the density solution. In many solutions, one layer or another is assumed to be "homogeneous", but the results, when plotted on the velocity–density diagram, are clearly inconsistent with this assumption; with the rejection of these, the range of self-consistent distributions is substantially narrowed. Most of the rejected solutions would be found unsatisfactory by other tests as well (Birch 1939, 1952) so that in general these tests are mutually consistent.

Figures 1 and 2 show the principal results of these comparisons. In Bullen's Model B (1950), the density is supposed to be determined by compressibility alone between depths of 80 km and 2 700 km. The velocity–density plot implies, however, a strong reduction of iron content with depth through the transition layer, thus an inconsistency with the premise, and this model is placed in the rejected list. There is additionally no evidence for the assumed discontinuity of density at the depth of 80 km.

Ramsey's (1949) distribution is essentially Bullen's Model A, modified by an increased density gradient in the lower mantle intended to signify an increasing iron content. This is consistent with the plot of the velocity–density diagram, which shows m increasing from about 23 at the base of Layer C to about 25 or so at the base of Layer D. From other evidence (Birch 1952, p. 261) it seems likely that Layer D is the most nearly homogeneous part of the mantle, and that so great a variation of iron content is inconsistent with the rate of variation of velocity with depth. Thus this distribution is also rejected.

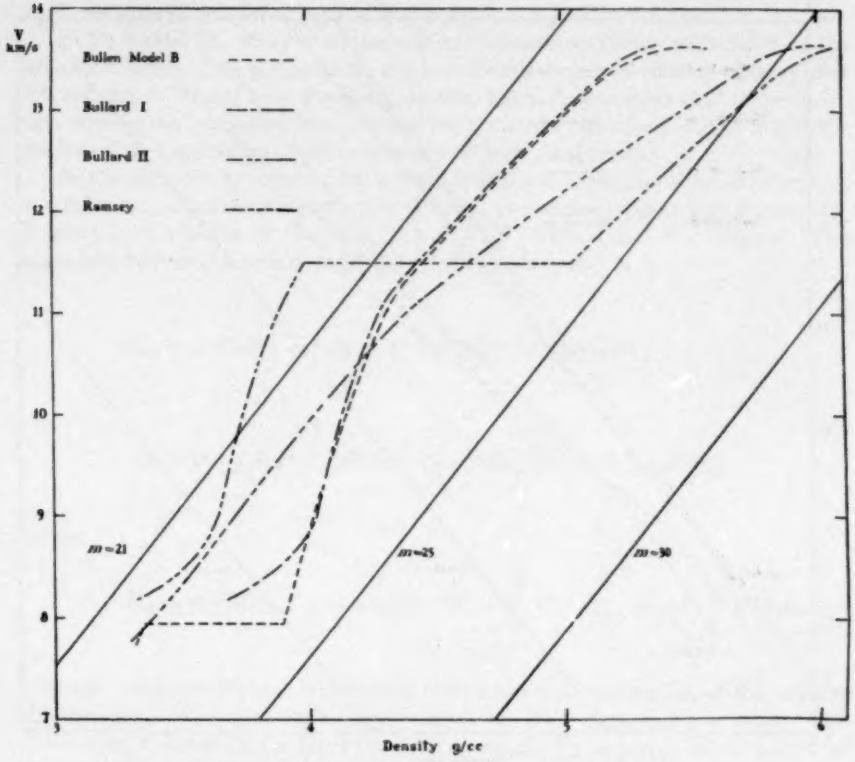


FIG. 1.—Velocity–density plot for rejected models.
Lines of constant mean atomic weight are indicated for m = 21, 25, 30.

Bullard's Model I (1957) is intended to be homogeneous throughout the mantle, that is, the change of density is everywhere determined by compressibility alone, as given by the seismic velocities. The velocity–density plot again shows a course inconsistent with the assumption: the mean atomic weight varies from about 24 in the upper mantle to 22 in the lower mantle. Thus the iron content varies, the resulting mantle is not homogeneous, and this model is rejected. In Bullard's Model II, a large discontinuity of density is introduced at a depth of 984 km, separating two layers in which homogeneity is implied by the use of the Adams–Williamson equation. The variation through the transition region is not consistent with the assumption of uniform iron content; furthermore, the plot

reaches regions to the left of the line for $m = 21$ which are probably excluded by considerations of elemental abundances. This model, like several of the others, was presumably not intended to have physical significance, as there is little encouragement from the seismic evidence for a major discontinuity at 984 km.

In Bullen's Model A (1953, pp. 217–218; recalculated as Model IV in Bullard 1957), the Adams–Williamson equation is applied above and below the transition layer, and in the core; in the transition layer, which had been shown to be inhomogeneous (Birch 1939, p. 477) the density was taken to be a quadratic function of

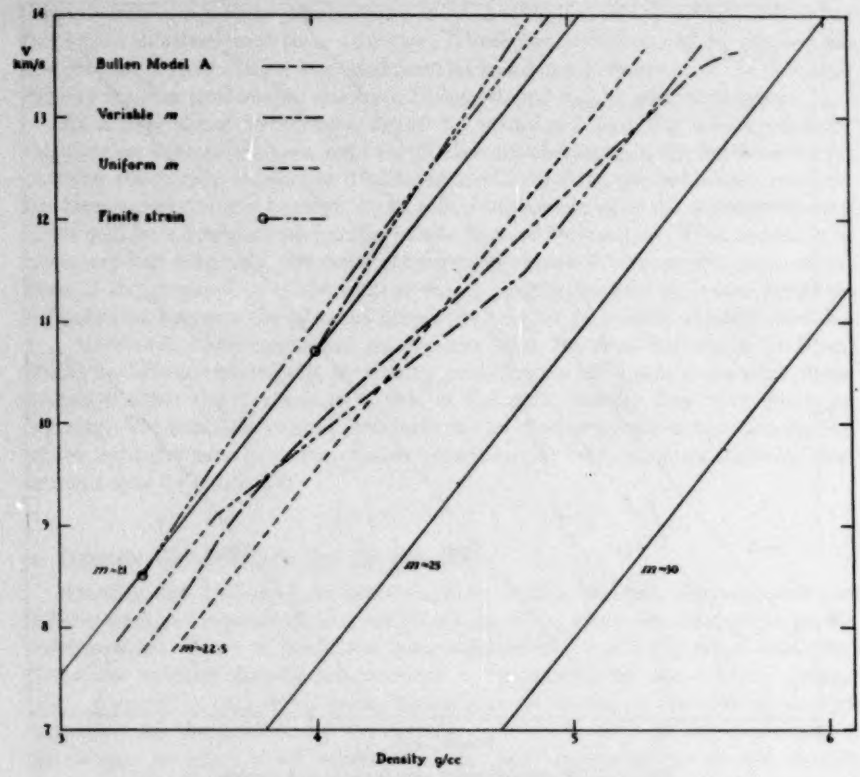


FIG. 2.—Velocity–density plot for self-consistent models.
Two calculated curves, according to the theory of finite strain, originate at the circled points.

radius, for interpolation between the densities at top and bottom and the density gradient at the base. This purely mathematical device results in densities which are very nearly proportional to the velocities throughout Layer C. The slope $dV/d\rho$ is 2·44 km/s per g/cm³ in Layer C, and nearly the same (2·34) in Layer D. The plot for this model on the velocity–density diagram consists of two nearly straight lines, showing first a decreasing iron content with depth in Layer B, then a trend toward higher iron content in Layers C and D. This distribution never departs greatly from the region of $m = 22$ to $m = 23$. The implication for the

transition region is that the increase of velocity is mainly the result of closer packing, greatly in excess of that arising from compressibility alone. We return to this point below. The difference in slope between these lines and the lines of constant m is now too small to lead to very sure conclusions as to departures from homogeneity. For reasons discussed below, some curvature of the lines of constant m may well take place at the higher velocities. On the other hand, this model is based on an adiabatic gradient of temperature; if the real gradient is higher, then the densities will rise less rapidly than in Model A. Bullard finds that a superadiabatic gradient of 2°/km results in a ratio $dV/d\rho$ for the lower mantle of 3·34 km/s per g/cm³, instead of 2·34 as in Model A.

In his Model III, Bullard adopts a linear increase of density with depth in the transition layer. This produces no marked discrepancy; the velocity–density plot follows that of Model A in the lower mantle, but indicates somewhat more variation of m in the transition layer, where the abnormal rise of velocity is implicitly attributed to a reduction of iron content with increasing depth.

As the simplest hypothesis, let us take m uniform throughout the mantle; this is equivalent, from the present point of view, to the assumption that the density is given by a relation of the form, $\rho = A+BV$, with A and B constants. The mass and moment of inertia of the mantle are then found as

$$M_m = 4\pi R_m^3 \int_c^1 \rho x^2\,dx = 4\pi R_m^3[A(1-c^3)/3+BW_1]$$

$$I_m = (8\pi/3)R_m^5 \int_c^1 \rho x^4\,dx = (8\pi/3)R_m^5[A(1-c^5)/5+BW_2]$$

where

$$R_m = 6338\,\text{km}, \quad c = 0{\cdot}548, \quad W_1 = \int_c^1 Vx^2\,dx, \quad W_2 = \int_c^1 Vx^4\,dx.$$

The two integrals W_1 and W_2 are found by numerical integration of the velocity distributions. With Jeffreys's (1939) velocities, $W_1 = 3{\cdot}097$, $W_2 = 2{\cdot}020$; with Gutenberg's velocities (as given in Bullard 1957), $W_1 = 3{\cdot}093$, $W_2 = 2{\cdot}017$, all in km/s. These differences are not significant for this problem. We proceed with the value of B given by the slope of the velocity–density relation, 0·302 g/cm³ per km/s. The undetermined constant A is then varied until the mass and moment of inertia of the mantle, and by difference, of the core, reach acceptable values. This calculation is greatly facilitated by Bullard's (1957) tabulation of Adams–Williamson solutions for the core; all these yield values of $I_c/M_cR_c^2$ close to 0·39. Two curves may be plotted with $I_c/M_cR_c^2$ as ordinate and M_c as abscissa: one from Bullard's Table 3, the other as found by variation of the constant A. The intersection determines the uniquely acceptable value of A to be 1·13 g/cm³; the values are shown in Table 4, and the corresponding densities for Jeffreys's and Gutenberg's velocities in Table 5. These density distributions satisfy the conditions on mean density and moment of inertia for the Earth as a whole when taken with the corresponding core.

Table 3

Estimates of parameters in velocity–density relations

$V = a(m) + 3{\cdot}31\rho$
$\rho = A(m) + 0{\cdot}302V$
V in km/s, ρ in g/cm³, m = mean atomic weight

m	a km/s	A g/cm³
21	−2·55	0·77
25	−5·7	1·72
30	(−8·6)	(2·6)

Table 4

Data for determining mass of mantle and core

Mantle of uniform m

A g/cm³	M_c 10^{27} g	$I_c/M_cR_c^2$
1·10	1·962	0·404
1·11	1·953	0·399
1·12	1·944	0·394
1·13	1·935	0·389
1·14	1·926	0·384
1·15	1·917	0·379

Adams–Williamson cores (Bullard 1957, Table 3)

Central density		
12	1·870	0·3897
13	2·007	0·3889

In order to find the value of mean atomic weight for this solution, it is necessary to interpolate between the lines of constant mean atomic weight. For $m = 21$, we have $\rho = 0{\cdot}77 + 0{\cdot}302\,V$; thus the density is increased uniformly by 0·36 g/cm³. This corresponds roughly to $m = 22{\cdot}5$. Random errors in the placement of the lines of constant m introduce an uncertainty in this value of at least 0·5 unit, and systematic errors cause it to be too high, as discussed below. Thus this solution is probably consistent with a mantle of uniform mean atomic weight resembling the average chondrite or the average mafic igneous rock (Table 2); it seems doubtful that the mantle can have so low a mean atomic weight as that of the Twin Sisters dunite or of the olivine nodules of basalts.

Figure 3 shows density as function of radius for Bullen's Model A, for the mantle of uniform mean atomic weight, and for a variant having the following properties: the density at 33 km is 3·30 g/cm³, Layers B and D are individually uniform, but with different values of m, the relation between velocity and density is linear in Layer C. Gutenberg's velocities have been used for the two new models. The values found for m, by interpolation, are for Layer B, 21·5, for Layer D, 23·4.

Table 5

Density distributions for the mantle

$x = R/R_m$ with $R_m = 6\,338$ km
Uniform m

	Jeffreys		Gutenberg*		Bullen Model A
x	V km/s	ρ g/cm³	V km/s	ρ g/cm³	ρ g/cm³
1·00	7·75	3·47	8·15	3·59	3·32
0·98	8·12	3·58	7·85	3·50	3·43
0·96	8·56	3·72	8·45	3·68	3·54
0·94	8·96	3·84	9·10	3·88	3·64
0·92	9·88	4·11	9·80	4·09	3·97
0·90	10·52	4·31	10·35	4·26	4·24
0·88	10·99	4·45	10·88	4·42	4·45
0·86	11·30	4·54	11·32	4·55	4·59
0·84	11·49	4·60	11·48	4·60	4·68
0·82	11·67	4·65	11·73	4·67	4·75
0·80	11·85	4·71	11·93	4·73	4·83
0·78	12·03	4·76	12·09	4·78	
0·76	12·20	4·81	12·26	4·83	
0·74	12·38	4·87	12·42	4·88	
0·72	12·54	4·92	12·57	4·93	
0·70	12·71	4·97	12·70	4·97	5·18
0·68	12·88	5·02	12·85	5·01	
0·66	13·01	5·06	12·98	5·05	
0·64	13·16	5·10	13·12	5·09	
0·62	13·32	5·15	13·26	5·14	
0·60	13·46	5·19	13·40	5·18	5·50
0·58	13·60	5·24	13·56	5·23	
0·56	13·64	5·25	13·70	5·27	
0·55	13·64	5·25	13·70	5·27	5·66

* See Bullard 1957; these velocities differ slightly from those of Gutenberg (1959) but agree substantially with those of the "Gutenberg case" in Dorman, Ewing & Oliver (1960, p. 92).

The velocity distributions now available for the mantle show no first-order discontinuities.* If there were density discontinuities, as in Bullard's Model II, there ought to be observable consequences for the reflexion of seismic waves, which do not appear to have been noticed. The restriction to continuous density distributions having a reasonable relationship to the velocities determines the density of the mantle within narrow limits. The distribution with uniform m represents a convenient norm. With lower densities (less iron) at shallow depths, higher densities are required in the deep mantle; Bullen's Model A probably comes close to the limit in this direction with the density at 33 km taken as 3·32. As an upper limit for the density at this depth, 3·6 might be generally accepted. This is hardly different from the value for the mantle of uniform m (Gutenberg's velocities). Thus these two distributions seem to be close to the allowable limits. It is, of course, understood that in speaking of these densities, mean values are intended; there is little evidence regarding the range of variation of density below the crust

* In one of Miss Lehmann's solutions, a "low-velocity" layer is separated by discontinuities from the distributions above and below; it is my understanding that this is a convenience of calculation rather than a physical requirement.

except that which may be deduced from the variation of velocity. If this is as much as 0·6 km/s, the density may range from 3·1 to 3·7 g/cm³. Like the velocities, these densities are means for large volumes.

The assumption of a linear relationship between velocity and density for large compressions leads to conclusions about uniformity consistent with other tests, and to a density distribution very similar to Bullen's Model A. This assump-

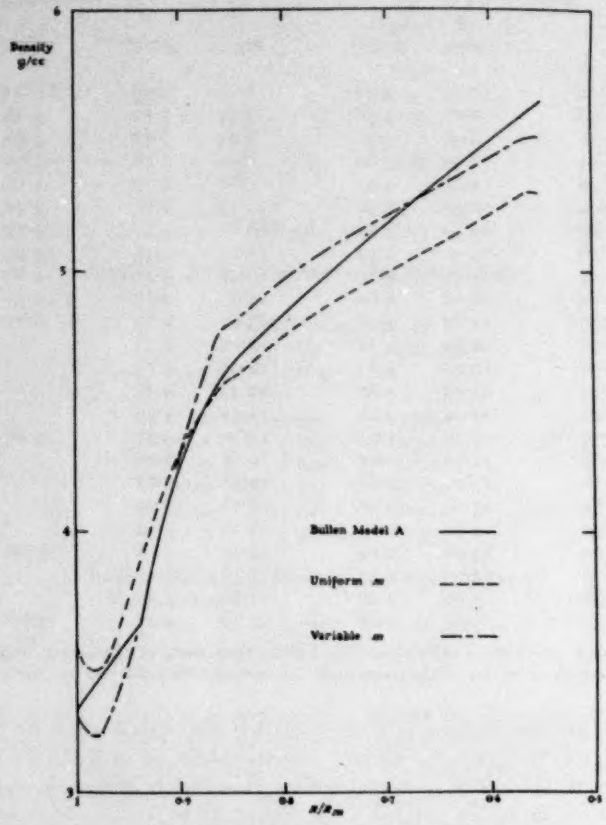


FIG. 3.—Density as function of radius for several models. R is radius, R_m the radius of the mantle, 6 338 km.

tion is not likely to be rigorously exact, and it would be helpful, pending further studies, to have an estimate of the probable deviations to be expected. The only investigation of rocks at very high pressures (Hughes & McQueen 1958) leads to apparently anomalous results. The density of a dunite seems to vary nearly linearly with pressure up to about 700 kilobars along the shock curve; this would mean a constant velocity over this whole range. Taken at face value, this study eliminates dunite as an important constituent of the lower mantle and transition layer. It seems likely that the application of shock-wave techniques to such materials has still to be perfected, and that reservations regarding these pioneering studies are warranted.

The general course of variation of velocity in the deeper mantle is consistent with an approximate result derived from the theory of finite strain (Birch 1939, p. 467). A velocity–density curve can be calculated by this method, starting at any point on the velocity–density diagram. For points on a line of constant m, this leads to a sheaf of curves which do not coincide, but have the general trend of the line. They diverge as the velocity increases, and at a velocity of 13 km/s, the spread of density is about 0·2 g/cm^3 (Figure 2). For these calculations, Poisson's ratio has been taken as 0·27 initially, and third-order elastic constants have been neglected. It seems possible that the typical curve will have a slight concavity toward the density axis.

Though not directly relevant, the shock-wave studies of the light metals indicate a nearly linear relation between "sound velocity" and density along the shock curves to pressures of the order of a megabar (Figure 5).

The most serious uncertainty is the effect of temperature. The experiments give an isothermal velocity–density relation; a rise of temperature decreases both velocity and density, but not in the ratio set by the slope of the isothermal compression curve. A specific effect of temperature is to be expected even at constant density, though this will be smaller than at constant pressure. In the application to the mantle, the temperature increases as depth and velocity increase; thus if the coefficients remained constant, points on the velocity–density lines should be corrected downward, to lower velocities, by increasingly large amounts as the density increases. This would have the effect of reducing the slope $dV/d\rho$ of lines of constant m, and of implying a higher density for a given V and m. This would bring the solution with uniform m closer to Bullen's Model A, and the inferred value for m would be reduced. At the same time, allowance for a superadiabatic gradient of temperature in the deep mantle increases the slope $dV/d\rho$ for the solutions by the generalized Adams–Williamson method, as shown by Bullard (Model I with temperature correction). The investigation of the effect of temperature remains as a future task. It seems likely that this will allow a reconciliation of all of the data with a mantle of nearly uniform mean atomic weight close to that of the average chondrite and a density distribution intermediate between Model A and the present solution for uniform m.

A detail common to the velocity distributions of Jeffreys and Gutenberg is the flattening of the velocity curve about 100 km above the core–mantle boundary. This is readily interpreted in terms of an increase of iron content with depth in this zone, as indicated in Figures 1 and 2. A closer examination of this feature might furnish information on diffusion in the deep mantle.

3. The nature of the transition layer

The transition layer, between depths of about 300 and 1 000 km, is distinguished by a rapid rise of velocity with depth, from about 8·6 to 11·5 km/s in 700 km; the increase of velocity below this is only 2·1 km/s in 2 000 km. The velocities in the transition region are in the range of the experimental velocities. Beginning with a material having an iron content similar to that of common ultramafic rocks, so great an increase of velocity requires a considerable increase of density as well, and simple compression alone, according to the elastic constants, is inadequate to account for the change (Birch 1939, p. 477; 1952, p. 259). An estimate of the increase of density to be attributed to phase change may be obtained from the new data.

The change of density in an interval of depth dz may be generalized to include the effect of change of iron content and of phase, as well as of ordinary compression and thermal expansion, as follows:

$$d\rho = g\rho\phi^{-1}(1-\alpha\phi\tau/g)\,dz+(\partial\rho/\partial m)_{T,P,y}\,dm+(\partial\rho/\partial y)_{T,P,m}\,dy.$$

The notation is that of an earlier paper (Birch 1952, p. 228), with the addition of m as mean atomic weight, and y a parameter which defines the closeness of packing or mean atomic volume. The "packing index" of Fairbairn (1943) might be

Table 6

Density changes on iron substitutions

Structure	Mineral	ρ	m	$\Delta\rho/\Delta m$	Ref.
Olivine	Fayalite	4·34	29·11	0·13	1
	Forsterite	3·21	20·10		
Orthopyroxene	Ferrosilite	3·96	26·39	0·12	1
	Enstatite	3·20	20·08		
Clinopyroxene	Hedenbergite	3·62	24·81	0·12	1
	Diopside	3·25	21·65		
Corundum	Corundum	3·98	20·39	0·11	2
	Hematite	5·26	31·94		
Spinel Aluminates	Hercynite	4·39	24·83	0·18	3
	Spinel	3·57	20·32		
Ferrites	Magnetite	5·15	33·08	0·13	3
	Magnesioferrite	4·58	25·58		
Chromites	Chromite	5·16	31·98	0·15	3
	Picrochromite	4·48	27·48		
Silicates	Fe_2SiO_4	4·85	29·11	0·14	4
	Mg_2SiO_4	3·55	20·10		
Garnet	Almandite	4·318	24·88	0·16	5
	Pyrope	3·582	20·15		

References for density

1. Tröger (1956).
2. Palache, Berman & Frondel (1955).
3. Gray (1957, p. 204).
4. Ringwood (1958a, 1958b). Dachille & Roy (1960) find 3·37 for the spinel of forsterite composition; both values are obtained by extrapolation
5. Skinner (1956).

used, but it is simpler to take the density itself; thus dy is the change of density arising from change of phase, at constant T, P, and m, in the interval dz. This equation is the generalization of the Adams–Williamson equation, to the extent that compositional variation may be represented by the single parameter m.

The various coefficients may be estimated from empirical data. Table 6 and Figure 4 show densities for various crystal structures as function of m. The most interesting substitution for present purposes is the magnesium–iron substitution. More information exists for the spinel structure than for any other (Gray 1957, p. 204), though most of the compositions are not especially relevant. The plot for spinels indicates the degree of scatter from a simple linear dependence of

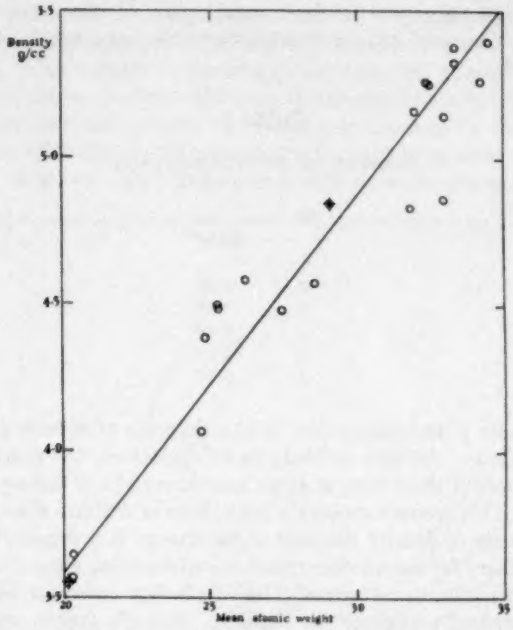


FIG. 4.—Density as function of mean atomic weight for spinels. The two points marked with crosses indicate Ringwood's values for the spinel forms of Mg_2SiO_4 and Fe_2SiO_4.

density upon m. The coefficient $(\partial\rho/\partial m)_{P,T,y}$ falls between 0·11 and 0·18 for all the magnesium-iron substitutions; the value 0·13 will be adopted for the following calculation.

The velocity has been supposed to depend only on ρ and m, so that

$$dV = (\partial V/\partial\rho)_m\, d\rho + (\partial V/\partial m)_\rho\, dm.$$

The coefficients are taken from Table 3: $(\partial V/\partial\rho)_m = 3{\cdot}31$ km/s per g/cm^3; $(\partial V/\partial m)_\rho = -0{\cdot}79$ km/s per unit of m. Combining with the equation for the density increment, we find

$$dV = (\partial V/\partial\rho)_m[g\rho\phi^{-1}(1-\alpha\phi\tau/g)\,dz + dy] + [(\partial V/\partial\rho)_m(\partial\rho/\partial m)_{T,P,y} + (\partial V/\partial m)_\rho]\,dm.$$

This may be integrated through the transition region with the aid of the seismic values of ϕ and the assumption that the various coefficients remain constant; also, the superadiabatic rise of temperature will be neglected. Between $R/R_m = 0{\cdot}94$ ($z = 413$ km) and $R/R_m = 0{\cdot}86$ ($z = 920$ km), the change of V is 2·32 km/s (Jeffreys 1939). We have

$$2{\cdot}32 = 1{\cdot}09 + 3{\cdot}31\Delta y - 0{\cdot}36\Delta m.$$

The first term on the right is the effect of ordinary adiabatic compression alone, less than half of the observed change. The coefficient of Δm is poorly determined, but a few values, shown below, illustrate the nature of the solution. In Bullen's Model A, the increase of density in the interval specified is 0·95 g/cm³, of which 0·33 g/cm³ is accounted for by ordinary adiabatic compression. The remaining 0·62 g/cm³ must be accounted for by a combination of phase change and change of iron content. The combination which satisfies the velocity relation is $\Delta m = 1{\cdot}0$, $\Delta y = 0{\cdot}48$ g/cm³.

Table 7

Solutions for the transition layer

Δm	Δy g/cm³
−1	0·26
0	0·37
1	0·48
2	0·59
3	0·70

If there were no phase change ($\Delta y = 0$) a decrease of m with depth would be required of 3·4 units. As m is unlikely to fall below 21, the material above the transition layer would then have a mean atomic weight of 24–25 and a density close to 4 g/cm³. This seems excessively high. If m is uniform through the transition layer, the change of density through phase change is 0·37 g/cm³. Estimates of the change of density for the olivine–spinel transformation range from $0{\cdot}35 \pm 0{\cdot}10$ g/cm³ (Ringwood 1958b) to 0·16 g/cm³ (Dachille & Roy 1960) for Mg_2SiO_4; Ringwood (1958a) obtained 0·51 g/cm³ for Fe_2SiO_4. Bernal's (1936) original estimate was 9 per cent, or about 0·3 g/cm³ for ordinary olivine. Thus it appears likely that there is little change of iron content throughout the transition layer, phase change accounting for nearly all of the anomalous increase of density.

4. Core

With the provisional solution for a mantle with uniform m, the parameters for the corresponding core are found by interpolation from Bullard's tables to be as follows: mean density, 11·03; density at a depth of 2 898 km, 9·96; central density (no allowance for the inner core) 12·47. In an earlier discussion (Birch 1952, p. 278) it was suggested that the density of the core exceeds that of iron at the same pressures by some 10 to 20 per cent. With the measurements now available of the density of iron under shock pressures equal to or greater than those of the core (Al'tshuler & others 1958; McQueen & Marsh 1960), this comparison may be made with more assurance; the density of iron at 1·3 Mb and about 2 000 °C is

11·2 g/cm³ and the difference with the above core is about 13 per cent at the core boundary. With Bullen's Model A, the difference is 15 per cent. A small amount of this may be removed on correction for melting and for core temperatures. The remaining difference may be accounted for by solution of lighter elements, perhaps silicon, as proposed by MacDonald & Knopoff (1958; see also Knopoff & MacDonald 1960) and Ringwood (1958c, p. 199; 1959). The hypothesis that the core consists of an iron alloy is now supported by closely relevant physical measurements as well as by the original geochemical arguments. The density of the inner core is not likely to exceed that of iron, about 13 g/cm³ at the corresponding pressure on the shock curve.

The evidence of the density–pressure relations may be supplemented by the velocity–density plot. Tables of the hydrodynamical "sound velocity", $(\partial P/\partial \rho)^{\frac{1}{2}}$, along the Hugoniot shock curve, are now available for large compressions for a number of metals (Rice, McQueen & Walsh 1958; Al'tshuler & others 1958; McQueen & Marsh 1960). This velocity is not exactly equal to the velocity along an adiabatic compression curve, but the difference is not significant for the present purpose. Seismology furnishes velocity of compressional waves for the Earth's liquid core and the two velocities for the mantle can be used to provide the comparable function $(K/\rho)^{\frac{1}{2}}$. Figure 5 shows these curves with those for the metals as far as tin

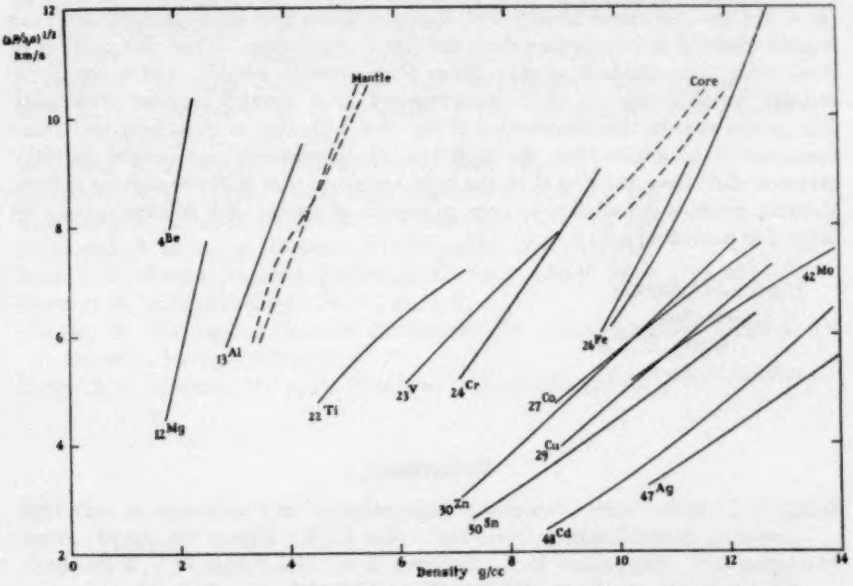


FIG. 5.—Hydrodynamical sound velocity as function of density for metals and for the Earth's mantle and core.

The velocities for the metals are from shock-wave studies to pressures of several megabars, principally by McQueen and Marsh. The second curve for iron represents the work of Al'tshuler & others. The numbers are atomic numbers.

(atomic number 50, atomic weight 118·70). The curves for the metals, nearly straight lines, are offset according to atomic number as far as the transition group; the curves for Zn, Cd and Sn, however, which do not crystallize in the cubic

system, overlap the curves for metals of lower atomic number, and the heavier metals are also irregular though their curves are beyond the relevant area. The geophysical interest of this diagram is that the curve for the core falls between those for Cr (24) and Fe (26), while the curve for the mantle lies in the light-element region, close to the curve for Al (13). Even without the information concerning chemical abundances, the conclusion that the mantle is composed primarily of light elements, the core of elements of the transition group, is strongly indicated.

This evidence is unfavourable to the hypothesis (Ramsey 1949) that the core is a metallic form of olivine or other silicate. The idea of a transition from a non-conducting to a metallic form was originally proposed for molecular hydrogen by Wigner & Huntington (1935); calculations by Kronig & others (1946) gave a transition pressure of 0·7 Mb with a density increase of about 100 per cent. Later work (Alder 1960; Behringer 1959) has indicated the uncertainty of such calculations. But in the application of this idea to the Earth's mantle, it should be remarked that solid molecular hydrogen is unlike the tightly-packed silicates and oxides of the light metals, which are denser than metals of the same mean atomic weight at ordinary pressure and even after compression to several megabars. For example, forsterite (m = 20·10) has the initial density 3·2; corundum (m = 20·39) the density 3·98, and so on, while the metal aluminium, intrinsically heavier (m = 27), has the initial density 2·7. Figure 5 shows that for a given velocity, the mantle material is also denser than the metal aluminium. Thus the packing in these ionic compounds is already closer than metallic packing, and a change to metallic form is unlikely to be accompanied by a marked increase of density. The proponents of this explanation of the core will have to show how the transformation to a metallic state can shift the velocity–density plot (or the density–pressure plot) from the region of the light metals to that of the transition metals. Reliable shock-wave studies to core pressures of oxides and silicates would be helpful in providing a final test.

Harvard University,
Cambridge,
Mass., U.S.A.:
1960 *September.*

References

Alder, B. J., 1960. State of matter at high pressure, in Conference on very high pressure, Botton Landing, New York. Also UCRL Report No. 5937T, 12 pp.

Al'tshuler, L. V., Krupnikov, K. K., Ledenev, B. N., Zhuchikhin, V. I., & Brazhnik, M. I., 1958a. *J. Exptl. Theoret. Phys.* (*U.S.S.R.*), **34**, 874. Translated in *Soviet Physics JETP*, **7**, 606 (1958).

Behringer, Robert E., 1959. *Phys. Rev.*, **113**, 787.

Bernal, J. D., 1936. *The Observatory*, **59**, 268.

Birch, Francis, 1939. *Bull. Seismol. Soc. Amer.*, **29**, 463.

Birch, Francis, 1952. *J. Geophys. Res.*, **57**, 227.

Birch, Francis, 1960. *J. Geophys. Res.*, **65**, 1083.

Bullard, E. C., 1957, *Verhandel. Nederlandsch Geol.-Mijnbouwkund. Genootschap.*, **18**, 23.

Bullen, K. E., 1950. *Mon. Not. R. Astr. Soc. Geophys. Suppl.*, **6**, 50.
Bullen, K. E., 1953. *Introduction to the Theory of Seismology* (University Press, Cambridge).
Clarke, F. W., 1924. U.S. Geol. Survey, Bull. **770**.
Dachille, Frank & Roy, Rustum, 1960. *Amer. J. Sci.*, **258**, 225.
Daly, R. A., 1933. *Igneous Rocks and the Depths of the Earth* (McGraw-Hill Book Co., New York).
Dorman, J., Ewing, M. & Oliver, J., 1960. *Bull. Seismol. Soc. Amer.*, **50**, 87.
Fairbairn, H. W., 1943. *Bull. Geol. Soc. Amer.*, **54**, 1305.
Gray, T. J. (Ed.), 1957. *The Defect Solid State* (Interscience Pub. Inc.).
Gutenberg, B., 1959. *Geophys. J.*, **2**, 348.
Hall, A. L., 1932. Mem. 28, Geol. Survey, Union of South Africa.
Hughes, D. S. & McQueen, R. G., 1958. *Trans. Am. Geophys. Union*, **39**, 959.
Jeffreys, Harold, 1939. *Mon. Not. R. Astr. Soc. Geophys. Suppl.*, **4**, 498.
Knopoff, L. & MacDonald, G. J. F., 1960. *Geophys. J.*, **3**, 68.
Lapadu-Hargues, P., 1953. *Bull. Soc. Geol. France*, **3**, 153.
MacDonald, G. J. F. & Knopoff, L., 1958. *Geophys. J.*, **1**, 284.
McQueen, R. G. & Marsh, S. P., 1960. *J. App. Phys.*, **31**, 1253.
Nockolds, S. R., 1954. *Bull. Geol. Soc. Amer.*, **65**, 1007.
Palache, C., Berman, H. & Frondel, C., 1955. *The System of Mineralogy*, Vol. I (John Wiley & Sons Inc., New York).
Poldervaart, A., 1955. Geol. Soc. Amer., Special Paper **62**.
Ramsey, W. H., 1949. *Mon. Not. R. Astr. Soc. Geophys. Suppl.*, **5**, 409.
Rice, M. H., McQueen, R. G. & Walsh, J. M., 1957. *Solid State Physics* 6, 1 (Academic Press, New York).
Ringwood, A. E., 1958a. *Bull. Geol. Soc. Amer.*, **69**, 129.
Ringwood, A. E., 1958b. *Geochim. et Cosmochim. Acta*, **15**, 303.
Ringwood, A. E., 1958c. *Geochim. et Cosmochim. Acta*, **15**, 195.
Ringwood, A. E., 1959. *Geochim. et Cosmochim. Acta*, **15**, 257.
Ross, C. S., Foster, M. D. & Myers, A. T., 1954. *Amer. Miner.*, **39**, 693.
Skinner, B. J., 1956. *Amer. Miner.*, **41**, 428.
Trŏger, W. E., 1956. *Optische Bestimmung der gesteinsbildenden Minerale* (E. Schweiz. Verlag, Stuttgart).
Urey, H. C. & Craig, H., 1953. *Geochim. et Cosmochim. Acta*, **4**, 36.

A Weak Layer in the Mantle?*

A. L. Hales

"You all do know this mantle."—*Shakespeare, Julius Caesar*, Act III, Sc. 2.

Summary

The paper discusses the effect of a weak layer in the upper mantle on the deformation of the crust arising from atmospheric loading, and on the isostatic compensation process. It also points out that *S* phases with periods of 1 second are much more strongly damped than *S* phases with periods of 10 seconds or more, and than *P* phases with short periods. The damping arising from a liquid layer is proportional to the square root of the frequency and does not provide a satisfactory explanation. Scattering of the kind suggested by Jeffreys, or any other mechanism leading to the firmoviscous law, gives rise to damping proportional to the square of the frequency and is sufficiently frequency sensitive to be consistent with the observations. The differences between *P* and *S* can only be accounted for in terms of different behaviour of the elastic constants for dilatational and distortional movements.

1. Introduction

Recent studies of the dispersion of the longer period surface waves at Lamont Geological Observatory and the Seismological Laboratory, Pasadena, have confirmed the existence of Gutenberg's low-velocity layer in the upper mantle (Dorman, Ewing & Oliver 1960). It was shown also that there are differences in the upper mantle structure below the oceans and below the continents. Kovach & Press (1960) have shown that there are differences in upper mantle structure within the Pacific Ocean basin. The low-velocity layer is characterized by a low rigidity and low shear velocity. In Dorman model 8099 the low-velocity layer extends over 160 km, beginning at a depth of 60 km below the oceans.

The confirmation of the existence of a low-velocity layer in the upper mantle in which the rigidity is low gives rise to the question whether the low-velocity layer is a world-encircling layer of low strength. This paper discusses some of the consequences of a layer of low strength.

2. Tilt Measurements

Observation of the non-tidal tilting of the Earth has been interpreted (Lettau 1937) as being due to changes in atmospheric loading of the crust and as implying

* Contribution No. 1004 from the Division of the Geological Sciences, California Institute of Technology.

an average rigidity of the crust of order $0{\cdot}3 \times 10^{11}$ dyn/cm^2. This value is much lower than the seismically inferred values of 3×10^{11} to 6×10^{11} dyn/cm^2 for the uppermost 200 km of the Earth. The inferences with regard to rigidity were made using a formula derived by Darwin (1907) for the deformation of the surface of a semi-infinite solid by a sinusoidal load. Darwin's equation is

$$w_1 = (g\sigma/2\mu k)\cos kx \qquad (2.1)$$

w_1 being the vertical component of the deformation of the surface due to the load $g\sigma \cos kx$. (It is assumed that the Earth is incompressible. A more direct derivation is given by Jeffreys (1959a).)

If there is a layer of near zero strength at depth then it can be shown that the deformations of the crust are much greater than in the case of a semi-infinite solid. Jeffreys (1959b) shows that for a load $g\sigma \cos kx$ on the surface of an elastic layer of thickness H ($H \ll 2\pi/k$) floating on a medium of density ρ^1, the deformation at the surface, w_2, is given by

$$w_2 = \frac{g\sigma \cos kx}{Dk^4 + g\rho^1} \qquad (2.2)$$

where

$$D = \frac{\mu(\lambda+\mu)}{3(\lambda+2\mu)}H^3.$$

It follows that

$$\frac{w_2}{w_1} = \frac{2\mu k}{Dk^4 + g\rho^1}. \qquad (2.3)$$

Values of this ratio for various values of H and $2\pi/k$ are given in Table 1.

Table 1

Ratio w_2/w_1

H, km \ $2\pi/k$, km	200	500	1 000	2 000
50	2·2	14·2	11·2	5·8
100	—	3·8	8·8	5·7
200	—	—	3·3	5·0

A weak layer at depth would result in larger deformations and larger tilts than predicted from the Darwin equation using the seismically determined rigidities and would thus be in accord with Lettau's observations. Tomaschek (1953) points out that the directions of the tilts resulting from atmospheric loading forces are not consistent with a direct effect on an elastic crust and suggests that non-elastic movements of individual tectonic blocks occur. It is essential to Tomaschek's hypothesis that there should be a weak layer in the upper mantle but the mechanical process involved is not sufficiently defined for quantitative comparison with observation.

3. The S phase

Ewing & Press (1956) pointed out that it is generally observed that the periods of S are many-fold larger than those of P and ascribed the absence of short-period components in S to the effects of transmission rather than of excitation. They suggested a mechanism of absorption and scattering involving the presence of pockets of magma in the mantle.

The author had noticed on records from the South African seismograph network that, in general, the ratio of S amplitude to P amplitude on the records from short-period Benioff vertical, or Willmore, seismographs was of the order of three for records at distances of 200 to 400 km, whereas at distances of more than 3 000 km the ratio was certainly less than $\frac{1}{4}$. At greater distances the ratio of S to P was even smaller. In fact, there is no evidence of an S phase on these records. For the same earthquakes, however, the Press–Ewing 15–80 at Pietermaritzburg, and the Benioff long-period vertical at Grahamstown show S phases with periods of 10 s, or more, which have two or three times the amplitude of the P phase.

Examination of the records of the Pasadena group of stations for shocks originating near the Pacific coast both north-west and south-east of Pasadena shows that the short-period S phase becomes too small to be recognized at distances of 600–1 000 km. The long-period S phase can often be found on records at distances of 1 000–2 000 km.

From the rough figures quoted it follows that the absorption of short-period S phases (1 s) is of the order of 12 times greater than that of the corresponding P phases between 300 and 3 000 km. If K_S and K_P are the coefficients of absorption for S and P waves then $\exp[-(K_S-K_P)x] = \frac{1}{12}$ for $x = 2\,700$. Gutenberg (1945) gives an average value of K of 0·12 per 1 000 km for longitudinal waves through the interior. It follows that $K_S = 1{\cdot}04$ per 1 000 km.

Supposing that there is a layer with viscosity ν, then for SH the equation to be satisfied by the velocity, v, is

$$\frac{\partial v}{\partial t} = \nu\frac{\partial^2 v}{\partial x^2}. \tag{3.1}$$

For

$$v \propto \exp(iwt) \,.\, \exp(i\alpha x)$$

$$iw = -\nu\alpha^2$$

and

$$\alpha = \sqrt{\left(\frac{w}{2\nu}\right)}(1+i). \tag{3.2}$$

The damping term is therefore $\{\exp[-\sqrt{(w/2\nu)}x]\}$ and the coefficient of absorption is proportional to the square root of the frequency. Damping within a liquid layer does not, therefore, account for the observed difference in the behaviour of the S phases with periods of 1 and 10 s.

Jeffreys (1959c) has shown that for elastico-viscous imperfections of elasticity the damping in a given distance is insensitive to the period. This is not in accord with the S phase observations.

Jeffreys (1959d) shows that the effect of scattering may be represented approximately by the firmoviscous law. For a transverse wave the displacement V must satisfy

$$\frac{\partial^2 V}{\partial t^2} = \beta^2\left(1+\tau\frac{\partial}{\partial t}\right)\frac{\partial^2 V}{\partial x^2}. \tag{3.3}$$

so that if $V \propto \exp(iwt)\exp(i\gamma x)$

$$\gamma \doteq \frac{w}{\beta}(1-\tfrac{1}{2}\tau i w) \tag{3.4}$$

and the damping term is $\exp[-\frac{1}{2}(\tau w^2/\beta)x]$. The difference in the coefficients of absorption for S phases of 1 and 10 s period would be 100 and would be sufficient to account for the observed differences. Jeffreys suggests, however, that the effect on longitudinal waves would be of the same order of magnitude as on shear waves. This is not the case.

Suppose, however, that the effect of internal friction is so much larger for distortion than for dilatation that the latter can be neglected. The displacement in a longitudinal wave will then satisfy the equation

$$\rho\frac{\partial^2 u}{\partial t^2} = \left(k+\frac{4}{3}\mu+\frac{4}{3}\mu\tau\frac{\partial}{\partial t}\right)\frac{\partial^2 u}{\partial x^2} \tag{3.5}$$

i.e.

$$\frac{\partial^2 u}{\partial t^2} = \alpha^2\left(1+\frac{4\beta^2}{3\alpha^2}\tau\frac{\partial}{\partial t}\right)\frac{\partial^2 u}{\partial x^2}$$

and if $u \propto \exp(iwt)\exp(i\gamma x)$

$$\gamma \doteq \frac{w}{\alpha}\left(1-\frac{2\beta^2 i w\tau}{3\alpha^2}\right). \tag{3.6}$$

The damping term is $\exp[(-2\beta^2\tau w^2/3\alpha^3)x]$ and the ratio K_S/K_P is $3\alpha^3/4\beta^3$, i.e. approximately 4. This is less than the value of $8\frac{1}{2}$ estimated from the observations. However, the value of K quoted from Gutenberg represents a mean value for longitudinal waves of all frequencies. It is probable that the value for 1 s waves should be greater than 0·12 per 1 000 km. Qualitatively the firmoviscous law satisfies the conditions: (1) that the damping of short-period S waves must be much greater than that of short-period P waves, (2) that the damping of short-period S waves must be much greater than that of long-period S waves.

The effective kinematic viscosity corresponding to the S wave coefficient of absorption is 10^8 poise.

Lomnitz (1956) suggested that the law of creep was logarithmic. For this law the absorption is proportional to frequency (Lomnitz 1957). Absorption varying as the frequency may be adequate to account for the difference between the amplitudes of S at periods of 1 and 10 s. It would, however, be necessary to postulate differences in the creep mechanism for P and S waves in order to account for the difference in absorption of short-period P and S waves.

4. Isostatic compensation

Jeffreys (1932) suggested that isostatic compensation took place in three stages. In the first stage, while the load is small, the stress-differences in the asthenosphere are less than the strength and the deformations and stresses are those for a surface load on a semi-infinite solid. When the load has increased to the point where the stress-differences at some depth exceed the strength, failure occurs and thereafter in the second stage, the stresses and deformations are those appropriate to a floating crust. In the third stage failure of the crust occurs and the final stage corresponds roughly to local compensation.

If there is a weak layer at depth then the first stage will be comparatively short. In the second stage the strong layer will be thicker than on the classical view in which the crust was the strong layer. The stresses and deformations corresponding to any given load will be small and greater departures from isostatic equilibrium will be possible before the crust fails. Thus the principal effect of a weak layer on the isostatic compensation process would be to introduce changes in the timing of the various stages.

It should be noted that if compensation takes place along the lines suggested by Jeffreys then it is necessary that there should be flow of material at some depth. Furthermore, movements of mantle material at some depth must be associated with each stage of the deformation of the crust. It may be that the flow is confined to limited regions in which there has been liquefaction as a result of plastic deformation as has been suggested elsewhere (Hales 1960).

Now let us consider the final result of the isostatic compensation process without regard to the steps by which it is achieved. It will now be assumed that compensation is local and detailed, i.e. that each column in the crust is in hydrostatic equilibrium. For the present purpose it is not necessary to take account of the curvature of the Earth and the isostatic equilibrium condition will be taken to be that the mass in all vertical columns is the same everywhere.

Suppose that there is a crust of thickness H_1, and uniform density ρ_1 over a layer of thickness H_2 in which the density is ρ_2; that below this layer the strength is zero and the density ρ_3; that a layer h of density ρ_1 is added at the surface and that as a result the columns descend a distance x.

It follows from the condition of equality of mass in adjacent columns that

$$\rho_1 H_1 + \rho_2 H_2 + \rho_3 x = \rho_1 (H_1 + h) + \rho_2 H_2 \qquad (4.1)$$

and therefore

$$x = h\rho_1/\rho_3. \qquad (4.2)$$

The small changes in density arising from the pressure and temperature differences at points x km apart have been neglected.

The height of the visible inequality at the surface is $h-x$, i.e. $h(1-\rho_1/\rho_3)$, and the root at the base of the crust is $h\rho_1/\rho_3$. It follows that the mass of the visible inequality is $h\rho_1(1-\rho_1/\rho_3)$ and the mass deficiency at the base of the crust is $(h\rho_1/\rho_3)(\rho_2-\rho_1)$ i.e. $h(\rho_1)(\rho_2/\rho_3-\rho_1/\rho_3)$. According to Bullen's distribution (Bullen 1953) the density at a depth of 100 km = 3·38, and at a depth of 200 km = 3·47, whereas $\rho_2 = 3{\cdot}32$. Taking $\rho_1 = 2{\cdot}67$ the mass of the visible inequality is $0{\cdot}56\,h$, and the mass deficiency at the base of the crust $0{\cdot}51\,h$, for $H_2 = 67$ km; for $H = 167$ km the corresponding figures are $0{\cdot}62\,h$ and $0{\cdot}50\,h$. It follows that although the compensation process originates in weakness at depth there is, nevertheless, a "root" at the base of the crust in which there is a deficiency of mass of the same order as that which would occur if the compensation originated at the base of the crust as in the classical theory of isostasy.

From the point of view of the gravity anomalies which result, it is instructive to consider the more detailed comparison of the densities in the loaded and unloaded columns given in Table 2. In this table z is the distance from the base of the crust in the "without additional load" column measured downward.

If there is no discontinuity in density at $z = H_2$ then the gravity anomaly (i.e. the Bouguer anomaly) arises wholly from the deficiency in density at the base of the crust.

If ρ_3 is greater than ρ_2, then there will be an additional negative anomaly from the deficiency in density at $z = H_2$. On the basis of present knowledge of the upper mantle it is unlikely that $\rho_3 - \rho_2$ will be as great as $\rho_2 - \rho_1$ and for this reason and the greater depth of the source the anomaly arising from the deficiency at $z = H_2$ will be very much smaller than that arising from the deficiency at the base of the crust.

Table 2

Comparison of densities in the columns, with and without the additional load

Range	Density in column with additional load	Density in column without additional load	Excess density in loaded column	Integrated effect over range interval
$-H_1 > z > -(H_1+h-x)$	ρ_1	0	ρ_1	$(h-x)\rho_1$
$0 > z > -H_1$	ρ_1	ρ_1	0	0
$x > z > 0$	ρ_1	ρ_2	$-(\rho_2-\rho_1)$	$-(\rho_2-\rho_1)x$
$H_2 > z > x$	ρ_2	ρ_2	0	0
$H_2+x > z > H_2$	ρ_2	ρ_3	$-(\rho_3-\rho_2)$	$-(\rho_3-\rho_2)x$

If ρ_3 is less than ρ_2, as might be the case if the weak layer is liquid, then there will be an additional positive anomaly superposed on the negative base of crust anomaly. As the density contrast due to a change from solid to liquid, is not likely to exceed 5 per cent, i.e. 0·17, and the anomaly has its origin at depth $z = H_2$ it will be much smaller than the negative anomaly arising from the deficiency at the base of the crust. A layer of low density at depth is, of course, unstable and can only be temporary.

This discussion has been carried out for ρ_2 a function of z instead of a constant and unless $d\rho/dz$ is large there is no significant difference in the results. It is important to notice that the equilibrium position is determined by the contrast between ρ_1, the density of the added material, and ρ_3, the density in the region where flow takes place, but that the gravity anomaly arises chiefly at the places where there are large density discontinuities. (As an example of this it should be noted that in the classical isostatic model the equilibrium position is determined by the contrast between the sub-Mohorovičić density and the added density, i.e. by densities of 3·3 and 2·7. If there is an intermediate layer, and the intermediate discontinuity and the Mohorovičić discontinuity move together, the gravity anomaly originates partly at the Conrad discontinuity and partly at the Mohorovičić discontinuity. The shape of the anomaly is, however, comparatively insensitive to the proportions originating in the two places, i.e. to how much of the density difference occurs at each discontinuity.)

There is one other case which is of interest. Suppose that in the weak layer the density increases fairly rapidly with depth, say $\rho_3 = \rho_4 + b(z - H_2)$, then (4.1) becomes

$$\rho_1 H_1 + \rho_2 H_2 + \int_{H_2}^{H_2+x} [\rho_4 + b(z-H_2)]\, dz = \rho_1(H_1+h_1) + \rho_2 H_2 \qquad (4.3)$$

and therefore

$$h\rho_1 = \rho_4 x + bx^2/2. \qquad (4.4)$$

The relation between x and h, and between $(h-x)$ and x is non-linear. Woollard (1959) shows such a non-linear relation between the Bouguer anomaly and the elevation of the topography, and between the depth to the Mohorovičić discontinuity and the elevation.

If y is the visible inequality, i.e. $h-x$, then from (4.4)

$$y = (\rho_1/\rho_1 - 1)x + bx^2/2\rho_1 \tag{4.5}$$

and

$$\frac{dy}{dx} = (\rho_1/\rho_1 - 1) + bx/\rho_1. \tag{4.6}$$

Woollard (1959) shows two curves. In his Figure 2 the Bouguer anomalies are plotted against surface elevation. In Figure 9 the surface elevation is plotted against the depth of the Mohorovičić discontinuity. From Figure 9 it is found that b is equal to $-0{\cdot}012\,\mathrm{g/cm^3km}$, whereas if it be assumed that the numerical value of the Bouguer anomaly is roughly proportional to x, Figure 2 would indicate a positive value for b. There are possibilities of systematic error in the determination of the depth to the Mohorovičić discontinuity by seismic methods and the scatter of the Bouguer anomalies is considerable. It would appear, however, that the non-linearity found by Woollard is due to some other cause.

The discussion in this section is concerned only with the changes in elevation resulting from post orogenic movements of the crust as a result of loading of the crust by deposition of sediments or erosion. It is not possible to extend the discussion to the orogenically produced inequalities of the surface because the mechanical processes involved there are more complicated and cannot be defined.

Since the strength of the crust is finite, the end result of the orogenic process will be a state approximating to isostatic equilibrium. Whatever the depth at which the flow takes place which makes possible the relative movement of adjacent columns, the gravity anomalies (Bouguer anomalies) associated with the compensation will originate at the places where there are large density differences, i.e. in the crust and at the Mohorovičić discontinuity.

5. Discussion

Of the methods discussed in this paper it would seem that only observations of the tilt or deformation of the Earth are likely to provide information with regard to the existence of a weak layer in the upper mantle.

6. Acknowledgments

This paper was written while the author was visiting the Seismological Laboratory, Pasadena. I have pleasure in acknowledging my indebtedness to Dr Frank Press, Director of the Laboratory, for much helpful discussion; to the California Institute of Technology, Dr Robert P. Sharp and Dr Press for the invitation to visit the Laboratory; and to the Carnegie Institution and the United States IGY Committee for the grant which made this visit possible.

Seismological Laboratory,
California Institute of Technology,
Pasadena, California:

1960 *October.*

(On leave from
Bernard Price Institute of
Geophysical Research,
University of Witwatersrand,
Johannesburg, South Africa.)

References

Bullen, K. E., 1953. *Introduction to the Theory of Seismology*, second edition (Cambridge).
Darwin, G. H., 1907. *Collected Papers*, **1**, 444. (Cambridge.)
Dorman, J., Ewing, M. & Oliver, J., 1960. *Bull. Seismol. Soc. Amer.*, **50**, 87.
Ewing, M. & Press, F., 1956. Abstract. *Trans. Amer. Geophys. Union*, **37**, 343.
Gutenberg, B., 1945. *Bull. Seismol. Soc. Amer.*, **35**, 57.
Hales, A. L., 1960. *J. Geophys. Res.*, **65**, 2155.
Jeffreys, H., 1932. *Mon. Not. R. Astr. Soc. Geophys. Suppl.*, **3**, 30.
Jeffreys, H., 1959. *The Earth* fourth edition. (Cambridge: (*a*) p. 199, (*b*) p. 203, (*c*) p. 254, (*d*) p. 41).
Kovach, R. & Press, F., 1960. In the press.
Lettau, H., 1937. *Z. Geophys.*, **13**, 25.
Lomnitz, C., 1956. *J. Geol.* **64**, 473.
Lomnitz, C., 1957. *J. Appl. Phys.*, **28**, 201.
Tomaschek, R., 1953. *Geofis. Pura e Applicata*, **25**, 17.
Woollard, G. P., 1959. *J. Geophys. Res.*, **64**, 1521.

Rock Magnetism in Western Europe as an Indication of Continental Growth

F. F. Evison

"Stones have been known to move."—*Shakespeare, Macbeth*, Act III, Sc. 4.

Summary

The palaeomagnetic interpretation of rock magnetism has led to an increasingly elaborate set of geodynamic postulates, which now include polar wandering, continental drift, and the rotation of continents and parts of continents. An alternative approach is suggested by the hypothesis of widespread continual plastic flow of basement rocks. Remanence data for western Europe are analysed from this viewpoint, assuming that the position of the poles has always been virtually the same as at present. The inferred pattern of flow is away from the high standing interior and towards the Northeastern Atlantic Basin. The amount of flow increases with the age of the rock; an accelerated rate of flow is indicated during the Hercynian revolution and a relatively slow rate in more recent times. These results are in accord with the concept of continental growth by plastic flow under gravity.

1. Introduction

Measurements of remanent magnetization have now been made on rocks in many different parts of the world, and from a wide range of geological periods. The rock samples collected from a given formation commonly show scattered directions of magnetization, but often, with adequate sampling, the mean direction for the formation can be estimated with a high degree of confidence. Results show that in many rocks older than about mid-Tertiary the mean direction of remanent magnetization is very different from the direction of the present geomagnetic field at the sampling locality.

The palaeomagnetic hypothesis has been widely adopted as a basis for interpreting these results. According to this hypothesis the direction of remanent magnetization in a magnetically stable rock is essentially the same as that of the ambient geomagnetic field at the time when the rock was formed. Certain complicating factors are recognized, such as subsequent heating or deformation. Hence metamorphic rocks are usually avoided, as also are known "thrust-sheets". On the other hand it is generally accepted that geological tilting can be corrected for in a straightforward manner, and one of the chief tests for magnetic stability depends on applying such corrections to the limbs of a fold.

The term palaeomagnetism, signifying ancient or fossil magnetism, can properly be applied to the remanent magnetization of rocks only if we are satisfied that the

ancient remanence vector has been effectively preserved. No one has established the general validity of this hypothesis.

There is much evidence from late Tertiary and Recent times that the geomagnetic field is closely similar, if we average out the secular variation, to that of a geocentric dipole aligned with the axis of rotation. On the assumption that this is true also for earlier times, the palaeomagnetic hypothesis implies that the remanent magnetic vector in a given rock formation indicates the position occupied by the geographic pole, relative to the sampling locality, at the relevant epoch. The main task of theoretical palaeomagnetism has been to find ways of reconciling the various pole positions given by rocks of the same age but situated at different places on the Earth's surface, and so to discover the actual locus that the pole is supposed to have traced out through geological time.

This notion of polar wandering has long been a matter of controversy among geologists and physicists but soon became an essential feature of the palaeomagnetic viewpoint. The polar locus derived from European rocks, however, was found to differ from that for North American rocks, and thus another controversial idea was revived—the hypothesis of continental drift—which, unlike polar wandering, had hitherto found little support among physicists. Jeffreys (1959) has shown that the hypothesis of continental drift is untenable on a number of physical grounds, among them that the rocks beneath the ocean floor must be expected to possess more rigidity than the continents. The picture of continents as rigid sheets has never been authenticated. In principle, rock magnetism could provide a critical test. For if coeval rocks at well-scattered localities in a particular continent were found to give consistent indications of the ancient pole position, and the corresponding pole similarly determined from another continent occupied a different position, the argument for rigid drifting continents would be strong. The remanence results so far available, however, are totally inadequate for such a purpose, and the important question of rigidity remains untested.

In practice it is doubtful whether the necessary continent-wide sampling of coeval rocks could be achieved, because sedimentation and igneous activity, in common with other geological processes, are apt to affect only limited regions at any one time. But to ascribe to an entire continent the results obtained from rocks at a single locality, or at most a few localities, is a considerable extrapolation.

In a very useful survey of published results on rock magnetism, Irving (1959) lists a world total of 112 estimated pole positions. Discussing the degree of similarity that has been found among results for rocks laid down in a given continent during a given geological period, Irving declares that "the agreement between the ten results . . . from Permian beds in five European countries is most striking". The sites from which these ten results were obtained, however, are included within an area of much less than continental extent, incorporating only about one-tenth of Europe. Nearly all European results for rocks of other ages are from sites in the same western European area. In other continents the sampling is much less adequate. The present paper will suggest a different approach to the interpretation of rock magnetism, in which observations will be related to conditions in the area from which they came, rather than to the continent as a whole.

Even the far-reaching postulates of polar wandering and continental drift have proved insufficient to account for certain more recent remanence observations. In order to preserve the palaeomagnetic hypothesis it has been necessary to add a third drastic proposal—that continents may rotate about vertical axes, in some cases through very large angles. Perhaps even more difficult to envisage is the suggested

independent rotation of portions of continents, such as part of the State of Oregon (Cox 1957; Irving 1959), Newfoundland together with parts of New Brunswick and Quebec (Nairn & others 1959; Du Bois 1959), and the Iberian Peninsula (Clegg & others 1957). In the case of Japan, it has been suggested on the one hand that a clockwise rotation has prevailed since the Miocene (Nagata & others 1959) and on the other hand that there has been an anticlockwise rotation during the past million years (Irving 1959).

The palaeomagnetic interpretation of remanence in rocks has thus called forth an increasingly elaborate set of geodynamic postulates. So many degrees of freedom have been invoked, considering the sparseness of the data, that the interpretation seems, if anything, less convincing now than formerly. The truth will only be established when a great many more observations can be taken into account. In the meantime there would seem to be grounds for seeking some alternative to the basic hypothesis of palaeomagnetism.

2. Plastic flow of basement rocks

The occurrence of plastic flow in basement rock as a widespread and long-continuing process has recently been proposed on geophysical, geological and physiographic grounds (Evison 1960). Actuated by gravity, the process bears some resemblance to the flow of ice sheets and glaciers; volcanic outpourings or orogenic uplift are analogous to snowfall in adding incrementally to the height of the sheet. The top few kilometres of the continental crust are believed to take part in the flow, which provides a mechanism for the expansion of continents by encroachment upon the surrounding oceans. This proposal of a slow plastic yielding of crustal rocks does not deny the clear evidence that the crust has great strength, for, as Jeffreys (1959, p. 210) has pointed out, the evidence in question is only concerned with the magnitude of stress that can be withstood without immediate fracture.

The flow process may also be supposed to operate more superficially and on a smaller scale within continents, where it seems to play a part in a variety of geological phenomena. As illustrations we may refer to two problems of the orogenic cycle that have been emphasized by Jeffreys (1959). The first of these is the mode of formation of the so-called "thrust-sheets" and nappes that form such a remarkable feature of many mountainous regions, stretching in some cases for hundreds of kilometres across the adjacent lowlands. Jeffreys (pp. 324–327) shows that this overlaying of rock upon rock is too extensive to be attributable to crustal shortening. He concludes from this that some superficial mechanism is required. Thrusting is no longer a possibility, and Jeffreys suggests sliding or flow down a slope (pp. 332–335). One may go further and rename the "thrust-sheet" an "overflow", for a sloping base is not actually necessary for plastic flow to occur; the overflow may even be due to plastic extrusion from the base of the mountains (Evison 1960).

Our second illustration arises from the problem of how the more ancient mountain ranges, such as the Scottish Highlands, came to be so reduced in height. As Jeffreys demonstrates (pp. 336–337), this cannot have been achieved by erosion alone; something like 10 km of rock would have to be eroded away to allow for isostatic adjustment. But old sediments are still present which could not have survived such denudation. Here again, plastic extrusion from the base of the mountains is a possible explanation. The lowering of mountains and the formation of nappes are thus both seen as results of the plastic flow of near-surface rocks under gravity.

Basement rocks in general are supposed to participate in the flow process, and it will be suggested below that this may have a substantial effect on remanent magnetization. Some writers on rock magnetism have specifically mentioned, however, that the rocks that they have studied are undistorted. Indeed it may be asked if the strains that must accompany large plastic flow and, on the present hypothesis, should be present in most basement rocks, could have gone unnoticed. One can only point out that the type of distortion to be expected in the process considered here is entirely different from that usually contemplated in geology, for the shear stresses involved are several orders smaller in magnitude (Evison 1960). Plastic flow has actually been invoked to account for a large rotation of the remanence vector in a folded Silurian limestone, as will be discussed more fully below; yet the usual petrological evidence of shear deformation was lacking (Graham 1949). Petrological studies are clearly required to elucidate this question.

Two macroscopic features of the flow of plastic sheets concern us here: the direction of flow and the effect on a vector embedded in the rock. Theory shows that the azimuthal direction of flow at any point in the sheet is that of the greatest downward slope of the free surface (Nye 1952). In a given region of a continent the general direction of flow is thus governed by the relative location of elevated areas and depressions. But since the process is a slow one the topographic features of dominant importance for the continental sheet are not the short-lived ones, even if very pronounced, but rather those that persist for many geological periods. Thus a major part is played by stable shields and ocean basins, and a minor one by mountain belts, inland basins, and geosynclines. "Thrust-sheets" are of course peculiarly a feature of the flanks of mountain ranges; they are of little importance in the present study, as it happens, but will doubtless need to be taken into account in other regions.

When a sheet of plastic material flows under gravity, the velocity of flow is greatest at the top of the sheet and diminishes steadily downwards. This has been established theoretically by Nye (1951) for flow near the equilibrium state, and has also been verified by observations on glaciers. For a thin horizontal sheet the movement is virtually confined to horizontal planes. Thus for any small volume the deformation is that of a simple shear. A short vertical vector that is embedded in the rock, in the sense that it may be identified with a specific set of particles, will be caused to tilt forward progressively in the direction of flow, whilst its length will increase as the secant of the angle of tilt. The effect on a non-vertical vector can be thought of as a similar tilting of the vertical component.

The flow hypothesis thus implies a progressive rotation at every point within a continental sheet or a "thrust-sheet". The proposed connection between flow and remanent magnetization is based on this rotational effect. Even with the continental sheet the effect of translation may be neglected since, as will be shown later, the amount contemplated is at the most a few kilometres per million years; and this movement pertains only to the plastic layer—the resulting continental expansion is only about one-tenth as fast, because of the down-building that must occur at the margin to preserve isostatic balance.

3. Flow and remanent magnetization

The remanent magnetization of grossly deformed rocks has been studied by several workers. In mica-schist and phyllite Daly (1959) has reported a strong

tendency for the remanence vector to lie in the plane of schistosity. Balsley & Buddington (1957) found that the direction of magnetization in a gneiss, containing magnetite as the predominant magnetic mineral, was closely related to the direction of lineation. In a haemoilmenite ore Carmichael (1959) demonstrated that the remanence was due to haematite lamellae which had exsolved parallel to the basal plane of the ilmenite in each grain. The ore displayed a preferred crystallographic orientation, and this had influenced the remanence vector, which tended to lie in the plane of the haematite lamellae (Hargraves 1959). In all these examples the authors suggest that stress has played some part in determining the direction of the remanence vector.

A number of rocks in which the magnetic susceptibility is anisotropic have been investigated by Howell & others (1958). In most cases these authors found that the remanence vector was nearly parallel to the plane of maximum susceptibility. Moreover, this plane was itself nearly parallel to the bedding in sedimentary rocks and to the foliation in metamorphic rocks. In these circumstances it seems quite unlikely that the direction of remanent magnetization should be that of the geomagnetic field when the rock was formed. Fuller (1960) has described lower Palaeozoic slates from Wales in which the cleavage was found to be a plane of high susceptibility with the remanence vector parallel to it. Where pyrrhotite was the chief magnetic mineral present, the susceptibility was sufficiently anisotropic to have defined the direction of magnetization, but in slates with haematite predominating it seemed that the remanence and the susceptibility had been aligned with the cleavage independently. Fuller concluded that the magnetic directions in the rock had been determined either directly or indirectly by the past stress environment. In these examples, indeed, the remanence vector has assumed the ultimate direction that gross deformation by shear can impart to a vector embedded in the rock.

Deformation by shear is a well-known accompaniment of folding. Graham (1949) has studied the direction of magnetization in a closely folded Silurian limestone. The directions observed on the two limbs of a particular fold were related in an entirely different manner from that in the more often quoted case of a Silurian sandstone, which was described by Graham in the same paper as an illustration of his stability criterion. In the limestone fold the remanence vectors for the rock in place were approximately parallel (Figure 1). Here again, according to Graham, plastic deformation at the time of folding appeared to have rotated the remanence vectors towards the bedding planes. The amount of rotation was actually much more than if the remanence vectors had simply been embedded in the rock; the magnetic shear, so to speak, was several times greater than the mechanical shear. This is illustrated in Figure 1.

When the direction of remanent magnetization is altered in this way by plastic flow associated with folding, the direction relative to axes outside the rock tends to be preserved (Figure 1). Further examples of this effect have been found by Kawai (Nagata 1953, p. 209). Doell's (1956) observations in steeply folded sandstone of Upper Miocene age might also be accounted for by the same mechanism. The magnetism in this rock was found to reside in a recrystallized mineral, and Doell has suggested that the remanence vector took the direction of the ambient field at the time of recrystallization. Since the proportion of recrystallized material was greater in the more intensely folded specimens, however, it seems possible that the remanence vector was reoriented essentially by shear, with recrystallization providing a mechanism for this change. The actual direction of magnetization

in this example was close to that of the present geomagnetic field and thus does not help us to choose between these explanations.

In published investigations involving both limbs of a fold it is rare to find that Graham's stability criterion, which allows for geological dip but not for flow, has been fully satisfied. Intensely folded Devonian sediments with anomalous directions of magnetization have been described by Creer (1957), who suggests that a thermo-remanent component may have been introduced at a subsequent reheating; it is conceivable, however, that plastic flow may have produced the observed rotations in this case also. In general it would appear unsafe to rely on remanence results obtained from steeply dipping beds, although such results, simply corrected for dip, have been widely used in palaeomagnetic discussions. It is a curious fact that the corrected direction in such cases is often rather close to the up-dip or down-dip direction, much as one would expect if the remanence vector had been reoriented by shear.

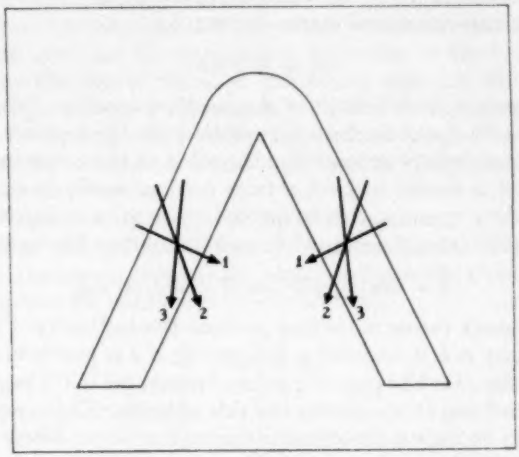


FIG. 1.—Effect of folding on a vector originally normal to bedding:
1. No shear.
2. Body vector rotated by simple shear.
3. Remanence vector rotated by shear.
(after Graham 1949).

Graham (1949) has pointed out two kinds of geological situation which allow the magnetic stability of rocks to be readily tested. The criteria thus formulated by Graham are often regarded as tests of the palaeomagnetic hypothesis itself, but this is quite unfounded. The first criterion is applicable when the rock in the formation being studied occurs also in a conglomerate; if the redeposited boulders or pebbles are magnetized in random directions one may conclude that the magnetism has not been subsequently affected by the ambient geomagnetic field. But this does not exclude the possibility that the direction of magnetization of the massive formation has been systematically altered, for example by plastic flow, any time before or after the conglomerate was formed.

The second criterion proposed by Graham states that if the directions of magnetization observed at various points in a fold are brought into parallelism by

correcting for the geological dip, then the magnetism is stable. Here too the direction of magnetization might well have been changed by plastic flow before folding occurred. Again, if the fold were subsequently incorporated in the continental plastic sheet or in a "thrust-sheet", some further flow could usually occur without seriously changing the angle between the magnetic vector and the bedding, both of which would undergo rotation in the same sense. These two field tests for magnetic stability in a rock can thus tell us very little, as Graham himself clearly realized, about whether the direction of magnetization has been altered by some widespread effect such as the flow processes considered here.

Since none of the geodynamic processes called forth by the palaeomagnetic hypothesis has found general acceptance on other grounds, the discussion which follows will assume that throughout geological history the geomagnetic field has always approximated to that of a geocentric dipole aligned with the present geographical poles, and that any discrepancy indicated by rock magnetism is due to plastic flow in the rock since it was formed.

Thus it will be assumed that at whatever epoch a rock was formed the original inclination I_0 of the remanence vector is given by

$$\tan I_0 = 2 \tan \lambda$$

where λ is the geographical latitude of the sampling locality. (Northern latitudes are counted as positive and southern as negative.) On the evidence given above for foliated and folded rocks it appears that the effect of shear deformation upon the remanence vector is similar in kind, though not necessarily in magnitude, to the effect of *simple shear* upon a vector embedded in the rock. Adopting this relationship one readily finds that the azimuth θ in which the flow has occurred is given by

$$\theta = \tan^{-1}\{\sin D/(\cos D - \tan I/\tan I_0)\}$$

where the remanence vector in its final position has inclination I and declination D. The ambiguity in θ is resolved as follows: (i) if I is positive (northern hemisphere), then θ and D are on opposite sides of north, (ii) if I is negative (southern hemisphere) then θ and D are on the same side of north. The amount of flow may be represented by an angle ψ analogous to the angle of shear, where

$$\psi = \tan^{-1}\{-\sin D/\tan I \sin \theta\} \qquad 0 \leqslant \psi \leqslant 90°.$$

The effect of continued flow upon any magnetic vector is ultimately:

$$I \to 0$$

and

$$D \to \theta + 180° \qquad I_0 > 0$$

or

$$D \to \theta \qquad I_0 < 0.$$

From the geometry of flow it can be seen that I retains the sign of I_0 throughout. A curious but indisputable feature of rock magnetism, however, is that the remanence vector in a given formation is sometimes found to be reversed. Western European rocks, for example, often show negative inclination. It is not yet known whether in general this is due to periodical reversals of the geomagnetic field or to some intrinsic property of magnetic rocks. Whatever the cause of reversal, we may use the equations in the preceding paragraph for these cases also, provided we take I

as positive and diminish D by 180 degrees. This gives the same results as if we were to rewrite the equations in the form appropriate to the conditions of reversal.

4. Direction of flow in western Europe

The quality of published results on rock magnetism varies greatly. Some results depend on so few observations that little reliance can be attached to them, considering the amount of scatter that is commonly found within a given rock formation. Here we shall concentrate on results that are based on 10 or more independent rock samples and show a well-defined grouping of values. The approximate age of the rock must of course be known. For reasons already stated we shall also confine ourselves to cases where the geological dip is small.

Certain other results are excluded because small errors in the estimated direction of magnetization could lead to large errors in the flow solution. This applies to studies in which the direction of magnetization is not far removed from that of the geocentric dipolar field and one cannot be sure that the secular variation has been averaged out. On this account most results for the Cenozoic will be omitted. Again, where the direction of remanence is very close to the horizontal a small error may change the sign of the mean inclination, and thus alter the calculated azimuth of flow by 180 degrees. Fortunately the Carboniferous and the Devonian in England can be represented by results in which, although the mean inclination is in each case only a few degrees from the horizontal, the values are supported by large numbers of observations.

Ten results from western Europe satisfy these requirements. They are detailed in Tables 1 and 2. The calculated flow for each case is depicted in Figure 2, in which an arrow centred on the sampling locality indicates the direction and amount of flow; the length of the arrow is proportional to ψ.

Table 1

Serial No.	Period	Formation	Locality	Reference
1	Miocene	Basic lavas	Skye	Khan 1960
2	Eocene	Antrim basalt	N. Ireland	Hospers & Charlesworth 1954
3	Jurassic	Radiolarite	Germany–Austria	Hargraves & Fischer 1959
4	Jurassic	Limestone	Germany–Austria	Hargraves & Fischer 1959
5	Triassic	Keuper sandstone	N. England	Clegg & others 1954
6	Permian	Lavas	Oslo	Rutten & others 1957
7	Permian	Lavas	S.E. France	Rutten & others 1957
8	Carboniferous	Dolerite sill	N. England	Creer & others 1959
9	Devonian	Sandstone	W. England	Creer 1957
10	Precambrian	Sandstone	N.W. Scotland	Irving & Runcorn 1957

The flow is directed, with one exception, towards the European margin of the Northeastern Atlantic Basin. This is in accord with the concept of flow under gravity. For on the one hand a permanently elevated hinterland is provided by the Baltic Shield—the only region of Europe to have remained consistently above sea-level since Precambrian times. To the south and east of this region lies the Russian

Table 2

(all angles in degrees)

Serial No.	Age (10^6 yr)*	No. of samples	ϕ	λ	D	I	θ	ψ	tan ψ
1	30	53	6 W	57 N	186	−60	194	14	0·26
2	60	24	6 W	55 N	194	−60	214	14	0·25
3	140	15	13 E	48 N	37	48	244	31	0·60
4	150	16	13 E	48 N	48	51	261	31	0·61
5	180	43	2 W	53 N	33	26	220	60	1·75
6	200	12	11 E	60 N	202	−33	207	52	1·28
7	220	14	7 E	44 N	217	−23	226	63	1·97
8	230	99	2 W	44 N	188	−5	188	85	11·09
9	300	35	3 W	52 N	196	−4	196	86	13·93
10	600?	81	6 W	58 N	123	44	315	51	1·23

Note.—In the computation of θ and ψ, negative values of I are taken as positive and the corresponding values of D are diminished by 180 degrees.

* Estimated on the scale given by Holmes (1944).

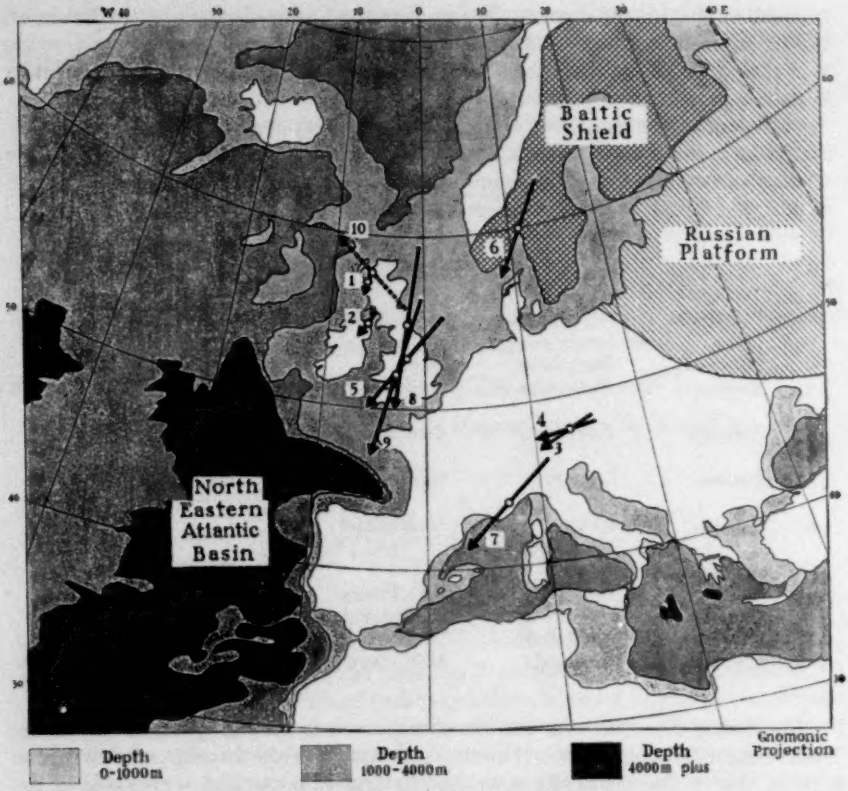


FIG. 2.—Rock flow in western Europe inferred from remanent magnetization.

Platform, where the shield has acquired a veneer of sediments but has never been far below sea-level. On the other hand the Northeastern Atlantic Basin constitutes a depression of great depth and extent, its permanency affected only by the gradual encroachment of the expanding continent. The region between shield and ocean basin is occupied by fold belts, where during long periods of sedimentation the basement rocks have lain well below sea-level.

The mobile boundary between continent and ocean is marked by the steep submarine slope that develops southwest of Ireland and extends into the Bay of Biscay, thence around the Iberian Peninsula to the coast of Morocco. No continental slope of comparable steepness is found anywhere off northern Europe, where the ocean floor lies at a much shallower level. The Mediterranean Sea is also relatively shallow and has not significantly influenced the pattern of flow in western Europe. Even on the Mediterranean coast of France the direction of flow since Permian times has been dominated by the Northeastern Atlantic Basin; this is consonant with the belief that the whole of the western Mediterranean was formerly above sea level (Klemme 1958).

The exception to the general pattern of flow is given by the Torridonian sandstone of northwest Scotland, which appears to have flowed in a northwesterly direction. This Precambrian rock is approximately twice as old as the next oldest in our group, and little is known about the configuration of continent and ocean during the first 200 million years or so after the rock was laid down. The inferred direction of flow may thus be the resultant of an ancient trend to the north, where the relatively small and shallow Norwegian Basin now lies, combined with the later southwesterly trend. We shall seldom have sufficient palaeogeographic data to be able to reconstruct regional surface gradients for Precambrian times. The complexity of the problem is illustrated by the fact that a number of samples from near the bottom of the Torridonian series gave a consistent direction of magnetization that was quite different from that of the bulk of the formation as considered here (Irving & Runcorn 1957).

An alternative interpretation of the Torridonian result is suggested by the close similarity between the inferred flow direction and the known direction of a remarkable series of "overthrusts" lying immediately to the east. These Palaeozoic movements, which include the famous Moine Thrust, have involved a relative displacement towards the west-north-west totalling at least 30 km (Anderson 1951, p. 112). Some outliers of Torridonian sandstone have actually been included in the recognized thrusts (Peach & others 1907, p. 6). It is attractive to suppose that the main Torridonian formation may have flowed outwards from the great Caledonian mountain chain at the same time as the "thrusts" were formed, though with less drastic petrological effects. This interpretation would imply, however, that circumstances had conspired to insulate the Torridonian from other movements, including the general continental expansion illustrated in Figure 2.

5. Amount of flow

A measure of the total flow achieved by a rock since its formation is given by the angle ψ which as we have seen is analogous to the angle of shear. As shown in Table 2 and Figure 2, the calculated values of ψ for western European rocks, with one exception, increase fairly steadily as the age of the rock increases. This is in accord with the proposal that the rotation is due to widespread and long-continued flow. Here too the exceptional case is that of the Precambrian sandstone of

north-west Scotland, which appears to have flowed only a small fraction of the amount that its great age would suggest. Thus both the direction and the amount of flow indicated for this formation lead to the conclusion that its flow history has been substantially different from that of the other examples.

A theoretical account of many aspects of flow has been given by Nye (1951), for the case of a long narrow sheet spreading laterally in response to an incremental increase of height. The two-dimensional flow of such a sheet is closer to the pattern that we have found in western Europe (Figure 2) than is the diverging flow of a circular sheet. The model is nevertheless a crude one and will serve only to indicate orders of magnitude.

Except at points near the edge of the sheet the velocity of the top of the sheet relative to the bottom is

$$v = 2\alpha(1 + h_0 r/h^2)$$

where r is the distance from the middle of the sheet,
h is the thickness of the sheet at the point in question,
h_0 is a constant for the material,
and α is the thickness of material added to the surface of the sheet in unit time. Within the limits of movement of a given rock formation the values of h and r may be taken as constant. Steady-state conditions are obtained by taking α as constant also. Then after time T the flow that has occurred may be represented by the angle of shear ψ', the mean value of which through the thickness of the sheet is given by

$$\begin{aligned}\tan\psi' &= vT/h \\ &= 2(\alpha T/h)(1 + h_0 r/h^2).\end{aligned}$$

This gives the angle of shear, assuming perfect plasticity in the material of the sheet. To the extent that rock is not perfectly plastic the actual angle will be larger (Nye 1957).

The following numerical values may be assigned so as to obtain an order of magnitude for $\tan\psi'$ (Evison 1960):

$$\alpha = 10^{-8}\,\text{km/yr}; \quad h = 2{\cdot}5\,\text{km}; \quad h_0 = 0{\cdot}00165\,\text{km}; \quad r = 1\,000\,\text{km}.$$

Here the value of α corresponds to a world rate of volume accretion sufficient to have produced all the continental crustal rocks now in existence, on the assumption that the production of this material, being a feature of the Earth's internal development, has proceeded at a constant rate by volume. As it happens, this value of α also corresponds to the estimated average rate of volcanic deposition on all land surfaces in historical times; the possibility is not excluded, however, that accretion may also occur at the bottom of the crust during orogenesis. The values of h and h_0 have been derived from a study of all the continents. The order of ψ' is then given by

$$\tan\psi' \mathrel{\widehat{=}} T \times 10^{-8}$$

where T is in years.

A given rock formation can flow only whilst it forms part of the plastic sheet, i.e. when its depth below the surface does not exceed about $2\frac{1}{2}$ km. In fold belts most rocks except the very youngest have been buried to much greater depths than this. Hence the actual relation between the angle of shear ψ' and the age of the

rock will be given by

$$\tan\psi' \triangleq aT \times 10^{-8}$$

where a is the fraction of its lifetime that the rock has spent in the plastic zone.

The angle ψ, which specifies the effect of flow upon the remanence vector, is analogous to the angle of shear ψ', but not necessarily of the same magnitude. The tendency in metamorphic rocks and some sedimentaries for the remanence vector to lie in the plane of foliation or bedding ($\psi = 90°$) suggests that the "magnetic shear" may be considerably greater than the mechanical shear, i.e.

$$\tan\psi > \tan\psi'.$$

This relationship is also indicated by Graham's (1949) observations of the magnetic effect of flow during folding. Hence to an order of magnitude we may write

$$\tan\psi = b\tan\psi' \triangleq abT \times 10^{-8} \qquad b \triangleq 10.$$

This equation gives the amount of "magnetic shear" that one would expect to find in western European rocks of various ages, if the rate of accretion α had been constant throughout the relevant time. The corresponding graph of $\tan\psi$ against age of rock is linear, and may be regarded as a characteristic of continental growth in the steady state. On the other hand the graph of observed values of $\tan\psi$ represents the actual course of growth, determined by a rate of accretion which rises and falls in synchronism with magmatic–orogenic activity.

The two graphs are shown in Figure 3, where the theoretical curve is drawn for $ab = 3$; this is tantamount to assuming that the average rate of growth has been about normal. A phase of rapid growth evidently accompanied the Hercynian

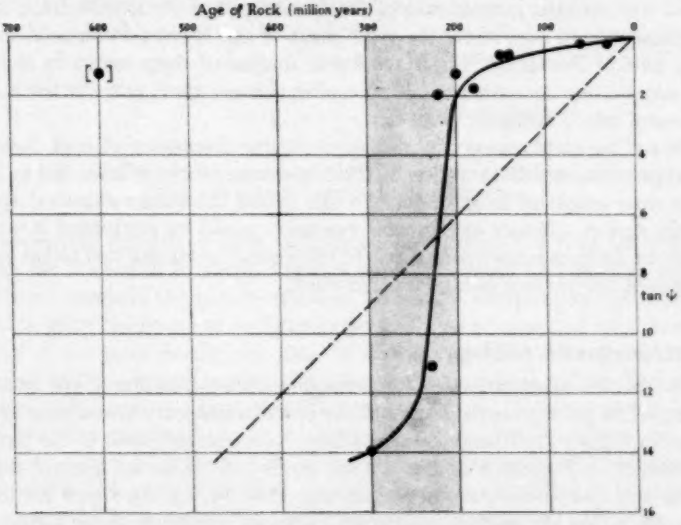


FIG. 3.—Growth of Europe inferred from remanent magnetization.

———— Actual growth curve.

– – – – – Hypothetical steady-state growth curve.

Shaded region indicates approximate duration of the main (Hercynian) period of magmatic–orogenic activity in western Europe (Umbgrove 1947)—evidently a period of very rapid growth.

revolution. For the last 150 million years, on the other hand, a slow rate of growth has prevailed. There is no evidence of accelerated flow associated with the Alpine revolution; in this connection it seems significant that much less magmatic activity was involved in the Alpine revolution than in the Hercynian (e.g. De Sitter 1956, pp. 471–474).

The total displacement of a rock due to flow may be estimated, to an order of magnitude, from the theory of Nye (1951). The mean horizontal velocity of the sheet at distance r from the middle is

$$v = \alpha r/h.$$

With the same numerical values as before this gives for steady state conditions

$$v \simeq 4 \times 10^{-6} \text{km/yr}.$$

The corresponding rate of expansion of a continent with crustal thickness 35 km is of order $vh/35$ or some 30 km in 100 million years. The distance travelled by a rock since its formation is avT. The value of a is likely to be about $\frac{1}{2}$, so that a rock travels on the average about 1 km for every million years of age.

Thus the Palaeozoic rocks of England must be supposed to have travelled several hundred kilometres since they were formed, and according to Figure 3 the greater part of the journey was achieved between 250 and 200 million years ago. If this picture is valid it has important implications for palaeogeography. Regions of sedimentary deposition, for example, must have been situated far to the north and east of where the resulting rocks now lie. Furthermore, at the onset of the Hercynian revolution one would have found the continental margin some 100 km landwards of its present position. The Northeastern Atlantic Basin would then have extended well into the present area of shelf southwest of the British Isles, and even into a present land area along the west coast of the Iberian Peninsula, including perhaps half of Portugal. The remarkable tongue of deep ocean in the Bay of Biscay may be seen as a remnant of an earlier oceanic zone, now for the most part transformed into continent.

It should be emphasized that this quantitative discussion of rock flow, continental expansion, and the rotation of the remanence vector is intended to indicate no more than orders of magnitude. To this extent the values obtained appear to be within reason. Before any greater precision could be attempted it would be necessary to develop a more sophisticated theoretical treatment and to have a better knowledge of the various parameters involved.

6. Palaeomagnetic analogy

When the ten observations of rock magnetism studied above are interpreted according to the palaeomagnetic hypothesis one obtains ten ancient pole positions as shown in Figure 4. The grouping of these poles corresponds to the pattern of flow exhibited in Figures 2 and 3. Thus poles 1–9 lie in the general region of East Asia and the northwest Pacific Ocean. Pole 10, on the other hand, is far removed from this region, the separation being attributed to polar wandering in the interval between about 600 and 300 million years ago.

The poles plotted in Figure 4 are usually taken to represent a simple curve, being the path relative to western Europe that the pole has traced out through time. There is some uncertainty as to whether the path should be drawn to pole 10a in East Africa or pole 10b in the central Pacific (Creer, Irving & Runcorn 1957); the flow hypothesis allows of no such ambiguity in the direction of flow.

The path from pole 1 to pole 9, allowing for the ovals of confidence around each pole, would actually seem to take an erratic course, unless there is some extraneous source of scatter. However, writers on palaeomagnetism have attached no special significance to deviations of this sort.

The trend towards lower latitudes in Figure 4 corresponds to the greater amount of flow achieved by the older rocks. The scatter in the latitude direction in Figure 4 corresponds to differences in the direction of flow in different parts of western Europe, and perhaps also to progressive changes in the flow direction as the continental margin advances on the ocean.

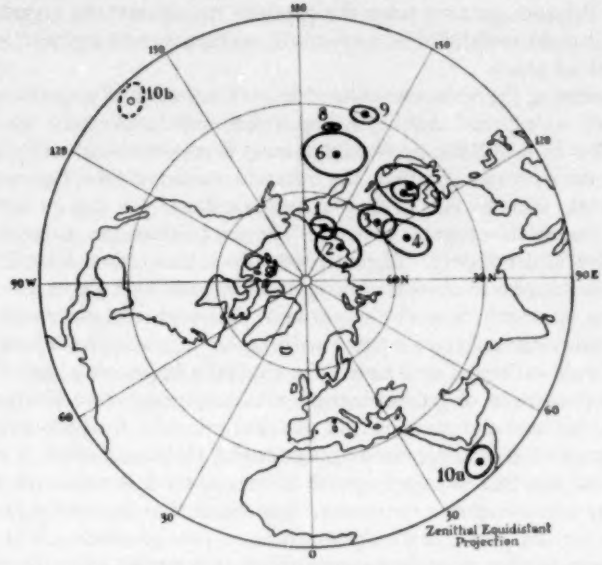


FIG. 4.—Pole positions according to palaeomagnetic hypothesis. (Pole 10b is in the southern hemisphere.)

An interesting feature of rock magnetism observations is that no systematic change in the direction of magnetization seems to occur for samples collected from different levels in the same formation. Since the sampling in some cases has embraced a great thickness of sediment one might have expected to detect some wandering of the pole during the time of deposition. The lack of evidence for this has led to the suggestion that the pole moves in jumps rather than steadily (Creer, Irving & Runcorn 1957). On the flow hypothesis, however, it is only to be expected that the effects of continental growth will be fairly uniform, as a rule, even through a very thick formation. To find a substantially different amount of flow one would look to an altogether different age of rock, while major differences in the direction of flow are only to be found in widely separated localities.

Observations of rock magnetism in western Europe do not of themselves bear upon the questions of continental drift or rotation. A comparison between the flow hypothesis and that of palaeomagnetism, relating specifically to these postulates, must await the analysis of observations from other parts of the world. This further analysis is at present somewhat hampered by the sparseness of data. It may be mentioned, however, that in some cases where the palaeomagnetic

interpretation appears to be least satisfactory a simple explanation is offered by the concept of continental growth. For example, the large clockwise rotation that has been suggested for the State of Oregon, U.S.A., since Eocene times, is replaced by a moderate amount of flow towards the Pacific Ocean.

7. Conclusions

The above study of rock magnetism in western Europe has been based on the view that the position of the Earth's rotational axis relative to the crust as a whole is likely to have greater permanence than the properties of any particular rock formation. Palaeomagnetism takes the opposite standpoint: the magnetic properties of certain rocks are held to be immutable, whilst a wandering habit is attributed to the terrestrial poles.

On interpreting the remanence observations for western Europe in terms of the hypothesis of widespread continual flow in basement rocks, one finds a simple pattern of flow in which the movement is away from the shield and platform and towards the ocean basin. Further, the inferred amount of flow increases with the age of the rock, and the rate of flow at a particular epoch can be related to the intensity of magmatic–orogenic activity. These inferences are in accord with the concept of flow under gravity. The interpretation is thus seen to be self-consistent.

These results seem to show that the palaeomagnetic hypothesis may not be the only possible approach to rock magnetism. Although the discussion has been limited to observations from a single small region, it is a region where rocks of a remarkably wide variety of ages have been studied. Again, only one of the major geodynamic postulates of palaeomagnetism is impugned, but this postulate is usually regarded as more plausible, on physical grounds, than the others are.

Any attempt to reconsider the implications of rock magnetism is made more difficult by the fact that so many reports of remanence measurements have taken an exclusively palaeomagnetic viewpoint. Sometimes the observed inclinations and declinations are not given, but only the inferred pole positions. The geological dip of the rocks is often described inadequately or not at all. Details of sampling procedure may be lacking. For reasons mentioned previously it has been thought desirable to restrict the present study to ten out of the much larger total of published results for Europe.

As with so many current geophysical problems, the final explanation of rock magnetism must await the amassing of many more data. In the search for an explanation the difficult and patient tasks of collecting rock samples from many lands, and determining the nature of the remanence vector, will always occupy the supremely important role.

8. Acknowledgments

The author wishes to thank Dr E. I. Robertson, Mr N. F. Barber, and Mr M. J. Randall for reading the manuscript and making many valuable suggestions. Discussions with numerous colleagues in the Geophysics Division have also been very helpful.

Department of Scientific and Industrial Research,
Wellington,
New Zealand:
1960 *August.*

References

Anderson, E. M., 1951. *The Dynamics of Faulting* (Oliver and Boyd).

Balsley, J. R. & Buddington, A. F., 1957. *Phil. Mag. Suppl.*, **6**, 317–322.

Carmichael, C. M., 1959. *Nature* (*Lond.*) **183**, 1239–1241.

Clegg, J. A., Almond, M. & Stubbs, P. H. S., 1954. *Phil. Mag.*, **45**, 583–598.

Clegg, J. A., Deutsch, E. R., Everitt, C. W. F. & Stubbs, P. H. S., 1957. *Adv. in Phys.*, **6**, 219–231.

Cox, A., 1957. *Nature* (*Lond.*) **179**, 685–686.

Creer, K. M., 1957. *Phil. Trans. Roy. Soc. A.*, **250**, 111–129.

Creer, K. M., Irving, E. & Runcorn, S. K., 1957. *Phil. Trans. Roy. Soc. A.*, **250**, 144–156.

Creer, K. M., Irving, E. & Nairn, A. E. M., 1959. *Geophys. J.*, **2**, 306–323.

Daly, L., 1959. *C.R. Acad. Sci. Paris*, **248**, 2614–2616.

De Sitter, L. U., 1956. *Structural Geology* (McGraw-Hill).

Doell, R. R., 1956. *Trans. Amer. Geophys. Un.*, **37**, 156–167.

Du Bois, P. M., 1959. *Nature* (*Lond.*) **184**, 63–64.

Evison, F. F., 1960. *Geophys. J.*, **3**, 155–190.

Fuller, M. D., 1960. *Nature* (*Lond.*) **186**, 791–792.

Graham, J. W., 1949. *J. Geophys. Res.*, **54**, 131–167.

Hargraves, R. B., 1959. *J. Geophys. Res.*, **64**, 1565–1578.

Hargraves, R. B. & Fischer, A. G., 1959. *Geophys. J.*, **2**, 34–41.

Holmes, A., 1944. *Principles of Physical Geology* (Nelson).

Hospers, J. & Charlesworth, H. A. K., 1954. *Mon. Not. R. Astr. Soc. Geophys. Suppl.*, **7**, 32–43.

Howell, L. G., Martinez, J. D. & Statham, E. H., 1958. *Geophysics*, **23**, 285–298.

Irving, E. & Runcorn, S. K., 1957. *Phil. Trans. Roy. Soc. A.*, **250**, 83–99.

Irving, E., 1959. *Geophys. J.*, **2**, 51–79.

Jeffreys, H., 1959. *The Earth* (Cambridge Univ. Press).

Klemme, H. D., 1958. *Bull. A.A.P.G.*, **42**, 477–512.

Khan, M. A., 1960. *Geophys. J.*, **3**, 45–62.

Nagata, T., 1953. *Rock Magnetism* (Maruzen).

Nagata, T., Akimoto, S., Skimizu, Y., Kobayashi, K. & Kuno, H., 1959. *Geophys. Notes*, **12**, Contribution No. 20.

Nairn, A. E. M., Frost, D. V. & Light, B. G., 1959. *Nature* (*Lond.*) **183**, 596–597.

Nye, J. F., 1951. *Proc. Roy. Soc. A.*, **207**, 554–572.

Nye, J. F., 1952. *Nature*, **169**, 529.

Nye, J. F., 1957. *Proc. Roy. Soc. A.*, **239**, 113–133.

Peach, B. N., Horne, J., Gunn, W., Clough, C. T. & Hinxman, L. W., 1907. *Mem. Geol. Surv. Gt. Brit.*

Rutten, M. G., van Everdingen, R. O. & Zijderveld, J. D. A., 1957. *Geologie en Mijnbouw*, **19**, 193–195.

Umbgrove, J. H. F., 1947. *The Pulse of the Earth* (Martinus Nijhoff).

VII. Atmosphere and Oceans

W. H. Munk & El Sayed M. Hassan, Atmospheric excitation of the Earth's wobble 339

K. Hidaka, Equatorial upwelling and sinking in a zonal ocean with lateral mixing 359

Atmospheric Excitation of the Earth's Wobble*

Walter Munk and El Sayed Mohamed Hassan

"'Tis here! 'Tis here! 'Tis gone!"—*Shakespeare, Hamlet*, I, 2.

Summary

Hassan (1960) has computed mean monthly values of the atmosphere's moments and products of inertia for the period 1873 to 1950, using all available station-level pressures. On the basis of these time series the excitation of the seasonal and 14-month (Chandler) wobble is discussed.

With respect to the seasonal wobble, our calculations confirm the conclusion by Jeffreys and others that it is due largely to atmospheric excitation. By working directly with station-level pressures we have avoided the procedure followed previously of removing the sea-level correction from sea-level pressure charts (thus "uncorrecting" for a large and undesirable correction). Our amplitudes are 25 per cent smaller than those given by Jeffreys (1959).

With respect to the Chandler wobble, it has been suggested by Jeffreys (1940), Rudnick (1956), and by Munk & MacDonald (1960) that this represents a resonance amplification of the irregular (non-seasonal) variation in atmospheric inertia. The computed spectral density of the atmospheric variation at the Chandler frequency falls short by one to two orders of magnitude to meet the requirement of this hypothesis. Excitation by irregular motion in the core is briefly considered. Here the electromagnetic coupling (limited by conductivity in the lower mantle) appears to be far too weak to account for the observed wobble. The excitation of the Chandler wobble must be considered an unsolved problem.

1. Introduction

A discussion of the Chandler wobble involves three principal topics: frequency, dissipation, generation. The observed frequency of 0·85 cycles per year (cpy) has been interpreted since the time of Kelvin and Newcomb as the Earth's Eulerian frequency of mutation (1·20 cpy) diminished by about 0·25 cpy due to the elastic yield of the Earth and by about 0·1 cpy due to the fluid yield of the oceans. The dissipation (damping time ~ 1 decade) has been attributed at one time or another to the Earth's fluid core, to the oceans, and to an elastic response of the mantle. The situation is confused. For a review we refer to Jeffreys (1959, Ch. 7) and to Munk & MacDonald (1960, Ch. 10).

One may regard the Chandler wobble as a resonance amplification of some (unknown) excitation; in this sense it is like playing a relatively broad-band noise

* This study was supported by a grant from the National Science Foundation.

through a moderately narrow pass-band filter. The filter characteristics (its peak frequency, f_0, and relative band width, Q^{-1}) can be discussed without reference to input, and this is essentially what has been done in the literature (Rudnick 1956; Fellgett 1961). In this paper, on the other hand, our concern is with the input problem.

It has been known since the investigation of Spitaler (1901) that the seasonal wobble (a spectral "line" at 1 cpy) is due largely to seasonal variations in the distribution of air mass; one is then tempted to ascribe the peak at 0·85 cpy to the *irregular* variation in the distribution of air mass. This was first done by Jeffreys (1940). Rudnick (1956) put the argument into the language of communications engineering: If the annual "carrier" is modulated in the sense that some winters (say) are more severe than others (amplitude-modulation) or earlier than others (frequency modulation), then sum- and difference-frequencies between modulation and carrier are generated. The difference frequencies form a band around the annual line. The Chandler wobble is then interpreted as a resonance amplification of the side-band energy at 0·85 cpy. Rudnick's argument is illustrated in Figure 1.

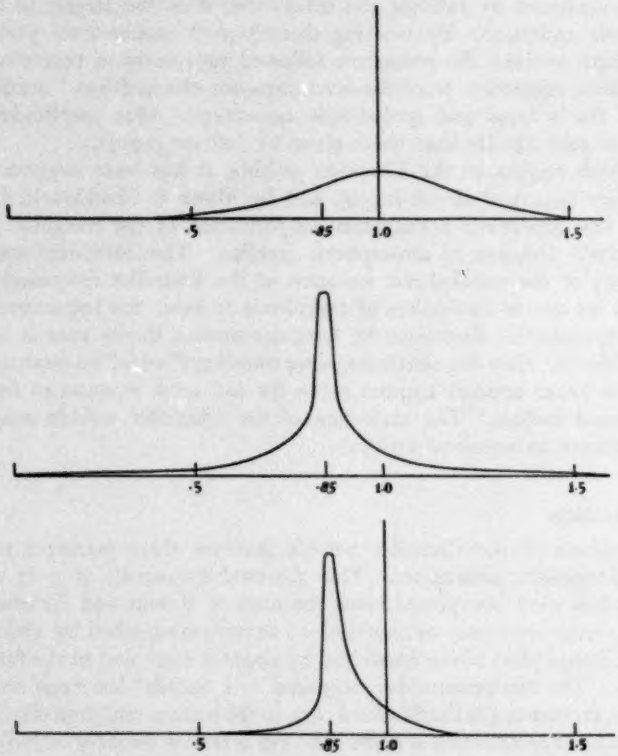


FIG. 1.—*Top:* the input spectrum is schematically presented by an annual line plus side bands. *Middle:* The power transmission of the Earth, peaked at the Chandler frequency, 0·85 cpy. *Bottom:* The power spectrum of the pole of rotation; this curve is the product of the upper two (adapted from Rudnick 1956).

The obvious procedure to test this hypothesis is to measure directly the atmospheric excitation spectrum at 0·85 cpy.

2. Atmospheric inertia

Let m_i, $(i = 1, 2)$, designate the angular displacements of the rotation pole toward Greenwich and 90° east of Greenwich, respectively. We wish to derive $m_i(t)$ arising from a variable distribution of surface load, $q(\theta, \lambda; t)\,\mathrm{g\,cm^{-2}}$, where θ is colatitude and λ east longitude. The equations of motion are

$$\psi_1 = m_1 - \dot{m}_2/\sigma_0, \qquad \psi_2 = m_2 + \dot{m}_1/\sigma_0, \tag{1}$$

where $\dot{m}_i = dm_i/dt$, and where

$$\begin{aligned} \psi_1(t) &= -\frac{a^4}{C-A}\int q(\theta,\lambda;t)\sin\theta\cos\theta\cos\lambda\,ds \\ \psi_2(t) &= -\frac{a^4}{C-A}\int q(\theta,\lambda;t)\sin\theta\cos\theta\sin\lambda\,ds \end{aligned} \tag{2}$$

designate the position of the *principal pole* (where the principal axis pierces the Earth's surface). Accordingly ψ_i equals the atmospheric products of inertia divided by $C-A$, where A, A, C are the Earth's principal moments of inertia. In the above equations

$$\sigma_0/2\pi = F_0 \tag{3}$$

designates the central frequency of the Chandler wobble (0·85 cpy), a the Earth's radius; $ds = \sin\theta\,d\theta\,d\lambda$ is an elementary area on the unit sphere.

Here $q(\theta, \lambda; t)$ is to be interpreted as the *departure* from the mean local load $\bar{q}(\theta, \lambda)$. This is derived from the departure, $p(\theta, \lambda; t)$, of the surface pressure from its local mean $\bar{p}(\theta, \lambda)$ according to $q = p/g$. Hassan (1960) has computed month by month values of $\psi_i(t)$ from published monthly pressures in accordance with equations (2) and subject to the following considerations:

(i) In deriving $\psi_i(t)$ there is no need to allow for the yield of the Earth under the superficial load. Equations (1), as written, allow for the cancellation between the effects of load deformation and rotational deformation (Munk & MacDonald 1960, Section 6.4).

(ii) There is observational and theoretical evidence that the ocean yields to the seasonable variation in air mass in the sense of an inverted barometer. The ocean then moves so as to annul horizontal pressure gradients on the sea bottom, but the bottom pressure varies in time because of the variable fraction of the atmosphere that is over the oceans.

(iii) The total mass of the atmosphere varies slightly on account of the variable content of water vapour, and we have prescribed an empirically determined seasonal variation in atmospheric mass.

(iv) With the yield of the oceans and the total mass of the atmosphere prescribed in accordance with (ii) and (iii), the equations (2) can be modified to give the excitation $\psi_i(t)$ in terms of continental observations only.

(v) The evaluation of $\psi_i(t)$ was done by fitting (in a least square sense) spherical surface harmonics to the observed values of $q(\theta, \lambda; t)$. Inasmuch as the stations are unevenly distributed, the orthogonality relations do not hold, and the amplitude of any one harmonic depends on the number of terms to which the expansion

is carried. In the present discussion all nine terms up to and including harmonics of degree 2 were evaluated.

All calculations were performed on an IBM 709 computer. Monthly values of $\psi_i(t)$ are tabulated in Hassan's paper.*

3. The observations

The monthly pressure values used in the calculations were taken from Volumes 29, 90, and 105 of the Smithsonian Miscellaneous Collections and from the World Weather Records published by the U.S. Weather Bureau. Figure 2 shows the position of *all* stations. Not all of these stations are available for any one year, and for the early part of the record only a small fraction is available. Figure 3 gives the number of stations used for any one year.

4. The seasonal variation

The seasonal variation is summarized in Tables 1 and 2. Two methods have been used for extracting the seasonal variation from the time series: (i) forming the average excitation, $\overline{\psi_i}(q)$, of all January values of ψ_i, then of all February values, etc.; (ii) computing the average load departure, $\bar{q}$, for all Januaries at each station and then computing the excitation $\psi_i(\bar{q})$ associated with the mean January departure, etc. Amplitudes of the annual and semi-annual terms are given in Table 2. $\odot$ is the longitude of the "mean Sun" measured from the beginning of the year (not from March 21). Allowance has been made for the fact that January values correspond to $\odot = 15°$ and not $\odot = 0°$, etc. The discrepancy between $\overline{\psi_i}(q)$ and $\psi_i(\bar{q})$ probably reflects the variation of station density during the time series. In 1900 there were 237 stations, mostly in Europe and North America. In 1950 there were 463 stations, and these were distributed more uniformly. We consider the values for $\psi_i(\bar{q})$ to give the best description of the atmospheric excitation.

Jeffreys's values (as adjusted by Rosenhead) are also given in Table 2. The largest coefficients are those in $\cos \odot$ for ψ_2, and in removing the unwanted sea-level correction the coefficient is reduced from $-25{\cdot}4 \times 10^{-8}$ to $-16{\cdot}3 \times 10^{-8}$. The magnitude of this adjustment points to the desirability of working directly with the uncorrected station-level pressure, as we have done. It is encouraging that Jeffreys's coefficients agree with ours as closely as they do.

For comparison we have included the excitation inferred in accordance with equation (1) from the astronomical observations of $m_i(t)$. The reduction by Jeffreys (1959) and by Walker & Young (1952) lead to results which differ considerably in phase. In all events the numerical values are roughly the same as those derived from the meteorological observations, and there is agreement concerning the predominance of ψ_2. A precise comparison is out of place because the effects of ground water and snow have not been considered. A rough estimate for these effects is (Munk & MacDonald 1960, Table 9.10)

$$\psi_1 = -0{\cdot}7 \cos \odot - 0{\cdot}1 \sin \odot, \qquad \psi_2 = 6{\cdot}4 \cos \odot + 5{\cdot}9 \sin \odot,$$

* R. Cowan has estimated the mean monthly atmospheric mass over four quadrants of the northern hemisphere for the period 1921 to 1930 (unpublished Master's Thesis, Mass. Inst. of Tech., 1950). The values are based on 80 monthly northern hemisphere station-level pressures. Cowan plots the difference in mass between opposite quadrants and finds these to bear a strong relation to the variable position of the pole.

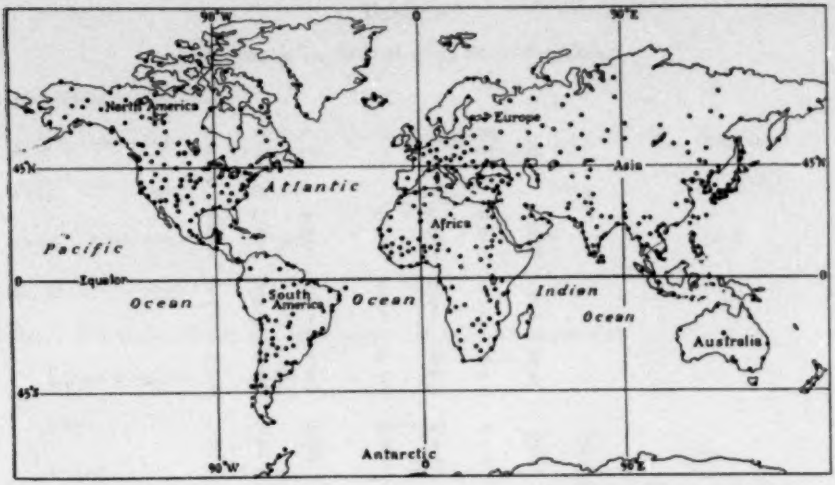


FIG. 2.—Position of stations used in the computation of excitation functions.

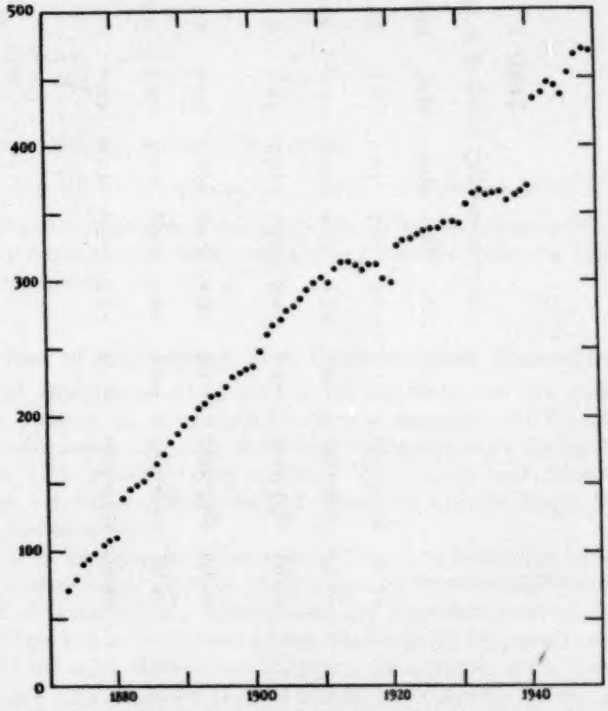


FIG. 3.—Number of stations used in the computation of excitation functions. The increments in 1921, 1931 and 1941 correspond to the publication of new summary volumes.

Table 1

Normal monthly values of ψ_1, ψ_2 in 10^{-8} radians

		Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
$\bar{\psi}_1$ (q)	1873–1950	−1·2	−0·5	+0·3	+0·5	−0·6	+0·5	+2·0	+1·2	−0·6	−0·6	−0·6	−0·8
$\bar{\psi}_1$ (q)	1900–1950	−0·8	+0·1	+0·2	+0·7	−0·8	+0·3	+2·1	+1·2	−0·8	−0·9	−0·3	−0·7
ψ_1 $(\bar{q})$	1873–1950	−1·6	−0·5	+0·1	+0·7	+0·1	+1·0	+3·3	+2·3	−0·4	−2·0	−1·3	−1·8
$\bar{\psi}_2$ (q)	1873–1950	−7·3	−6·0	−3·9	+0·3	+3·3	+7·9	+10·8	+8·4	+2·1	−3·2	−5·4	−7·3
$\bar{\psi}_2$ (q)	1900–1950	−8·8	−6·9	−4·8	+0·4	+4·2	+9·7	+12·8	+10·1	+2·4	−4·1	−6·5	−8·8
ψ_2 $(\bar{q})$	1873–1950	−11·1	−8·3	−5·4	+1·4	+6·5	+12·2	+15·4	+12·1	+2·7	−5·8	−8·1	−11·4

Table 2

Harmonic coefficients of ψ_i in 10^{-8} radians

Meteorological Observations					
$\bar{\psi}_1(q)$	1900–1950	$-0{\cdot}7\cos\odot$	$+0{\cdot}2\sin\odot$	$+0{\cdot}2\cos 2\odot$	$+0{\cdot}5\sin 2\odot$
$\psi_1(\bar{q})$	1873–1950	$-1{\cdot}8$	$+0{\cdot}2$	$+0{\cdot}4$	$+0{\cdot}8$
$\bar{\psi}_2(q)$	1900–1950	$-10{\cdot}2$	$-1{\cdot}4$	$+1{\cdot}7$	$+0{\cdot}8$
$\psi_2(\bar{q})$	1873–1950	$-12{\cdot}9$	$-1{\cdot}0$	$+1{\cdot}8$	$+1{\cdot}4$
Jeffreys–Rosenhead (before and after removal of sea-level correction)					
	before correction	$-1{\cdot}6$	$-1{\cdot}6$		
ψ_1	after	$-1{\cdot}7$	$-0{\cdot}9$		
	before	$-25{\cdot}4$	$-6{\cdot}8$		
ψ_2	after	$-16{\cdot}3$	$-1{\cdot}6$		
Astronomical Observations					
ψ_1	Jeffreys	$-0{\cdot}5$	$-0{\cdot}5$		
ψ_2	Jeffreys	$-15{\cdot}5$	$+6{\cdot}8$		
ψ_1	Walker & Young	$-4{\cdot}8$	$+6{\cdot}3$		
ψ_2	Walker & Young	$-7{\cdot}3$	$+15{\cdot}0$		

which together with our values for $\psi_i(\bar{q})$ gives

$$\psi_1 = -2{\cdot}5\cos\odot + 0{\cdot}1\sin\odot, \qquad \psi_2 = -6{\cdot}5\cos\odot + 4{\cdot}9\sin\odot$$

for the geophysical estimates of the excitation. The good agreement with astronomical observations is now destroyed, and one wonders about the validity of the hydrological estimates.

5. Comparison of meteorological and astronomical time series

The most straightforward procedure for inquiring into the causes of the wobble is to compare the astronomically inferred excitation with the meteorologically derived excitation. Figure 4 shows the two functions for the last 12 years of observations. The result is disappointing. The curves bear no resemblance. Furthermore the astronomically derived values show much larger fluctuations (note difference in scale).

The seasonal variation has been removed from both time series by subtracting from each January value the mean of all Januaries, from each February value the mean of all Februaries, etc. Furthermore the operation involved in applying equation (1) acts as a rejection filter against the Chandler frequencies which are so prominent in the $m_i(t)$ observations (compare, for example, with Munk & MacDonald, Figure 7.4). Equation (1) as it stands does not allow for any dissipation. To allow for dissipation, the terms

$$-\tfrac{1}{2}Q^{-1}m_2, \qquad +\tfrac{1}{2}Q^{-1}m_1 \tag{4}$$

must be added on the right sides of equations (1). Here Q^{-1} is the usual dimensionless description of the width of the spectral peak, and the assumption of no dissipation corresponds to $Q = \infty$. Taking $Q = 30$ for a reasonable estimate, we have calculated the astronomically inferred excitation and find it not to differ significantly from the curve plotted in Figure 4 for $Q = \infty$.

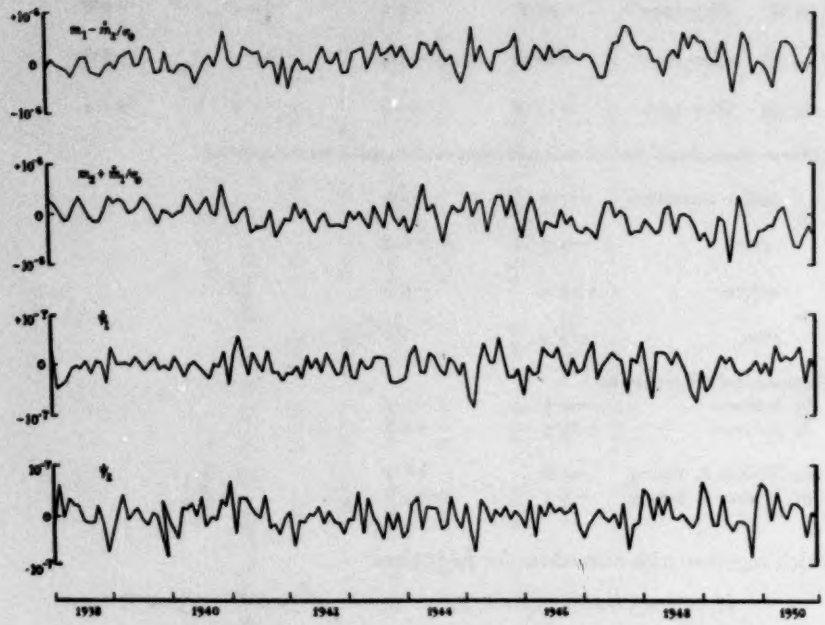


FIG. 4.—Components of the non-seasonal excitation function inferred from the astronomical observations (upper curves), and derived from meteorological observations (lower curves). The units are in radians; the scale for the meteorological excitation is magnified tenfold relative to astronomically inferred excitation. The position of January means is indicated on the time axis.

6. The spectra

Figure 5 shows the spectra and all possible cross-spectra between the astronomical and meteorological time series, after removal of the seasonal terms. The spectra are to be interpreted as the contribution, per unit frequency band (in cpy), toward the mean square value of the record (in radians^2). The ordinates represent then a "power density", and the units are radians^2/cpy. The central and bottom parts of the figure give coherence, R, and phase, ϕ, to be interpreted as follows: if any two records, $x(t)$ and $y(t)$ are played through identical narrow frequency filters centred at some frequency f, and the filtered $x(t)$ is lagged relative to the filtered $y(t)$ by various amounts until maximum correlation is achieved, then $R(f)$ is this maximum correlation and $\phi(f)$ the corresponding phase lag. The calculation is actually performed digitally by computing auto- and cross-correlations and forming Fourier transforms. The computing parameters are summarized in Table 3. For further details we must refer to an account by Blackman & Tukey

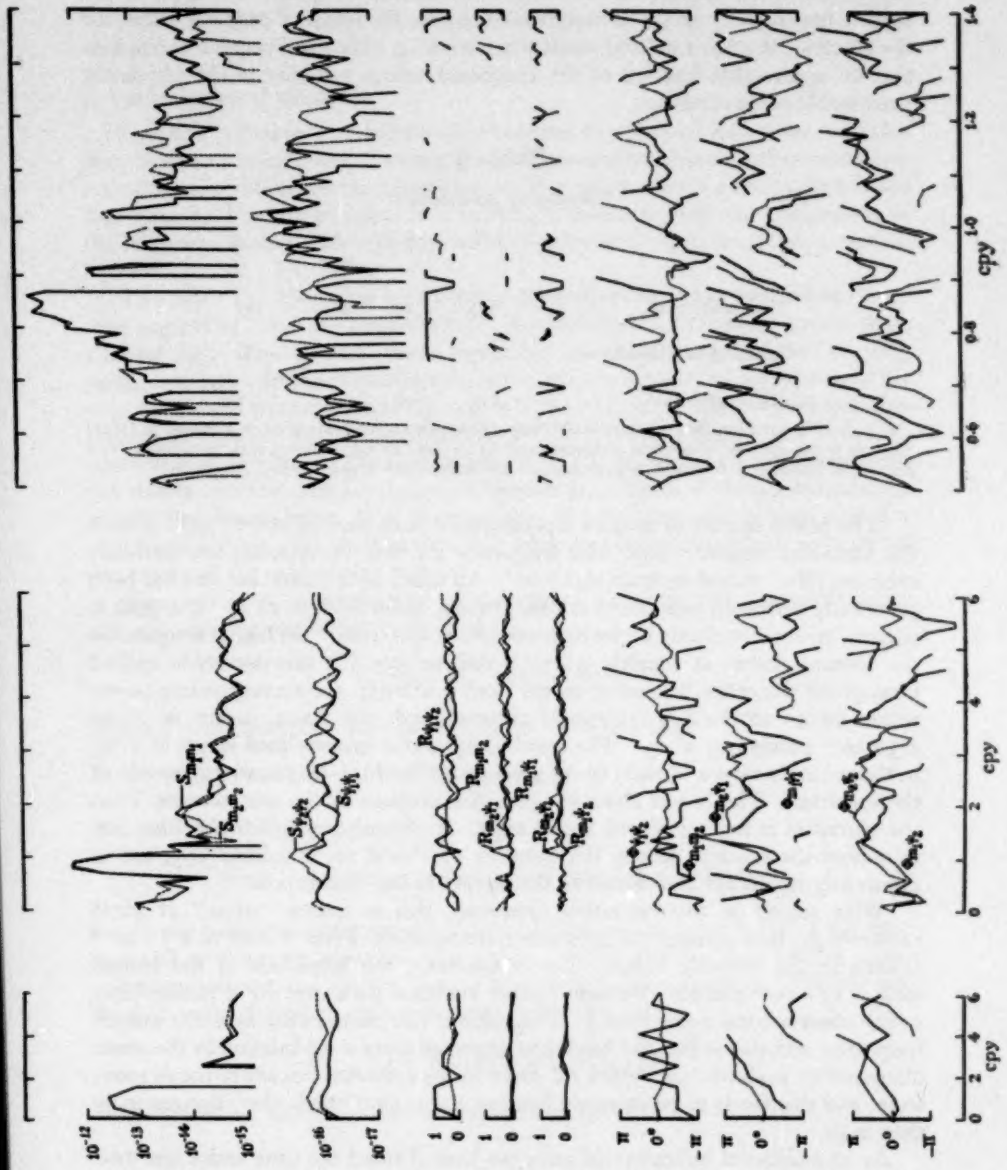


FIG. 5.—Spectra and cross-spectra of the astronomical and meteorological observations. The three displays correspond (from left to right) to increasing frequency resolution, at the expense of an increasing uncertainty in the spectral estimates. The upper curves show the power spectra in radians²/cpy); the central curves give coherence, and the lower curves phase relations between all pairs of time series.

(1958) and to Appendix A.2 of Munk & MacDonald (1960). The gaps in the spectra (particularly severe at high resolution) are the result of negative values for the spectra. It is the expected result when working with small degrees of freedom that an appreciable fraction of the computed values are outside the physically permissible range of values.

Table 3

*Computing parameters**

	N	m	Δf cpy	ν
Low resolution	613	12	0·5	101
Intermediate resolution	613	60	0·1	20
High resolution	613	600	0·01	1·5

* N is the total number of values in the time series, taken at intervals of Δt = 1 month = (1/12) year. m is the number of spectral estimates, and $\Delta f = (2m\,\Delta t)^{-1}$ is the frequency resolution. The degrees of freedom, $\nu = (2N/m) - \frac{1}{2}$, determine the uncertainty in the spectral estimates.

The power spectra of m show a pronounced peak centred at 0·85 cpy. This is the Chandler wobble. Near this frequency the two components are perfectly coherent ($R = 1$) and m_1 leads m_2 by 90°. All this is as it should be, and has been previously discussed (see Munk & MacDonald, 1960, Section 10.1). The peak is narrow, and resolved only in the high resolution spectrum.* At higher frequencies the spectra flatten at roughly 4×10^{-15} radians²/cpy.† If this density is applied throughout the entire frequency range, from 0 to 6 cpy, the corresponding power would be $4 \times 10^{-15} \times 6 \approx 2{\cdot}4 \times 10^{-14}$ radians², and the r.m.s. value is about $1{\cdot}5 \times 10^{-7}$ radians, or 0″·03. The conclusions is that uncorrelated errors of 0″·03 in the monthly values of $m_i(t)$ could account for the high-frequency asymptote of the spectrum. This is just about the precision claimed by the astronomers. Thus the spectrum is not significant above 2 cpy. At frequencies below the Chandler frequency the spectral density lies between 10^{-14} and 10^{-13} radians²/cpy, and is apparently significant and related to the motion of the "mean pole".

With regard to the excitation spectrum, this is nearly "white" at 10^{-16} radians²/cpy, thus corresponding to uncorrelated errors of $(10^{-16} \times 6)^{\frac{1}{2}} \approx 2{\cdot}5 \times 10^{-8}$ radians in the monthly values. For comparison, the amplitude of the annual term is 13×10^{-8} radians. Without further evidence we accept 10^{-16} radians²/cpy as the observational noise level.‡ A significant rise occurs only near the annual frequency, and this is just the behaviour expected from a modulation in the sense discussed by Rudnick (see Figure 1). Here the ψ_2 spectrum lies above the ψ_1 spectrum, and this too is to be expected because the annual line is also stronger in ψ_2 than in ψ_1.

As an additional indication of error we have divided the time series into two portions. The dividing line at 1920 was selected on the basis of the station-time

* The negative values to either side of a peak are of no physical significance; for a sharp peak the Tukey method gives negative side bands between $2\Delta f$ and $3\Delta f$ to either side of the peak.

† The high-frequency spectrum shows weak but significant peaks centred at 1·7 and 1·84 cpy (outside the range of Figure 5). These correspond to (2 × 0·85) cpy and (0·85 + 1) cpy and thus raise the intriguing possibility of a nonlinear interaction between the Chandler peak and the annual line.

‡ The coherence between ψ_1 and ψ_2 at higher frequencies may be interpreted as evidence to the contrary.

plot (Figure 3). We expected a lower noise level for the modern observations, but this is not the case.

7. Cross-spectral analysis

Equation (1) expresses a relationship between two (vector) time series, and this then leads to a comparison between the astronomically inferred and meteorologically determined excitations. Figure 4 is an illustration of such a comparison in the *time domain*. For our purposes it is more convenient to make the comparison in the *frequency domain*. Accordingly we shall transform equation (1) as a relation between spectra.

The following discussion represents a generalization of the method of "multiple regression" and of "path analysis" in statistics (for example, Bartlett 1956, Chapter 14). The problem is to determine the relative contribution of $y_1(t)$, $y_2(t)$, ..., to $x(t)$. In the classical formulation one assumes $x(t) = a_1y_1(t)+a_2y_2(t)+$ $+$... as if there were no phase lag and as if the admittance did not vary with frequency. These are impossible restriction to geophysical problems, and the cross-spectral analysis removes the restrictions. Unfortunately the needed formulae are not readily available (see Kendall 1956, Section 9.3). Some of the needed relations will be developed below. The discussion is sketchy, particularly with regard to significance levels. We shall have to rely on future investigations for an adequate discussion of the numerical results.

First we need to summarize a few rules. Consider any two time series, $x(t)$, $y(t)$, and let correlation be defined by

$$R_{xy}(\tau) = \langle x(t)y(t-\tau)\rangle$$

where $\langle\ \rangle$ designates an average over time. It follows that

$$R_{xy}(\tau) = R_{yx}(-\tau).$$

Now let $\dot{R} = dR/d\tau$ and $\dot{x} = dx/dt$. Then

$$\dot{R}_{xy}(\tau) = R_{\dot{x}y}(\tau) = -R_{x\dot{y}}(\tau) = +R_{\dot{y}\dot{x}}(-\tau) = -R_{\dot{y}x}(-\tau).$$

In the present exercise we assume no friction and no noise in either the astronomical or the meteorological time series. Equations (1) then predict certain relations between the spectra plotted in Figure 5. For example, multiply $m_1(t)$ by $\psi_1(t-\tau)$ and average each term with respect to time:

$$R_{m_1\psi_1}(\tau) = R_{m_1m_1}(\tau)+\sigma_0^{-1}\dot{R}_{m_1m_2}(\tau). \tag{5}$$

The spectra are related to the correlations according to

$$R_{xy}(\tau) = \int_0^\infty [S_{xy}(f)\cos 2\pi f\tau + T_{xy}(f)\sin 2\pi f\tau]\,df \tag{6}$$

where $S_{xy}(f)$ is termed the co-spectrum and $T_{xy}(f)$ the quadrature spectrum. It follows that

$$S_{xy}(f) = S_{yx}(f), \qquad T_{xy}(f) = -T_{yx}(f).$$

For ordinary power spectra, $x = y$ and hence $T_{xy}(f) = T_{yx}(f) = 0$. Note that

$$\sigma_0^{-1}\dot{R}_{xy}(\tau) = \int_0^\infty r[-S_{xy}(f)\sin 2\pi f\tau + T_{xy}(f)\cos 2\pi f\tau]\,df$$

where

$$r = \sigma/\sigma_0 = f/f_0.$$

Substituting (6) into (5) we obtain a relation between three integrals, and this is certainly satisfied if it is satisfied at each frequency. Equating cosine and sine terms leads to the equation in the first line of the following summary formulae. The others have been derived in a similar manner. The dependence of all spectra on f is to be understood.

$$\left.\begin{aligned}
&S_{m_1\psi_1} = S_{m_1m_1} + rT_{m_1m_1} && T_{m_1\psi_1} = -rS_{m_1m_2}\\
&S_{m_1\psi_2} = S_{m_1m_2} && T_{m_1\psi_2} = T_{m_1m_2} + rS_{m_1m_1}\\
&S_{m_2\psi_1} = S_{m_1m_2} && T_{m_2\psi_1} = -T_{m_1m_2} - rS_{m_2m_2}\\
&S_{m_2\psi_2} = S_{m_2m_2} + rT_{m_1m_2} && T_{m_2\psi_2} = rS_{m_1m_2}\\
&S_{\psi_1\psi_2} = |1-r^2|S_{m_1m_2} && T_{\psi_1\psi_2} = (1+r^2)T_{m_1m_2} + \\
& && \quad + r(S_{m_1m_1} + S_{m_2m_2})\\
&S_{\psi_1\psi_1} = S_{m_1m_1} + 2rT_{m_1m_2} + r^2S_{m_2m_2} && S_{\psi_2\psi_2} = S_{m_2m_2} + 2rT_{m_1m_2} + r^2S_{m_1m_1}.
\end{aligned}\right\} \quad (6)$$

It is sometimes more convenient to express all results as functions of the excitation spectrum. By proper elimination of terms between equations (6) we find

$$\left.\begin{aligned}
&S_{m_1\psi_1} = \frac{S_{\psi_1\psi_1} - rT_{\psi_1\psi_2}}{|1-r^2|} && T_{m_1\psi_1} = \frac{-rS_{\psi_1\psi_2}}{|1-r^2|}\\
&S_{m_1\psi_2} = \frac{S_{\psi_1\psi_2}}{|1-r^2|} && T_{m_1\psi_2} = \frac{-rS_{\psi_2\psi_2} + T_{\psi_1\psi_2}}{|1-r^2|}\\
&S_{m_2\psi_1} = \frac{S_{\psi_1\psi_2}}{|1-r^2|} && T_{m_2\psi_1} = \frac{rS_{\psi_1\psi_1} - T_{\psi_1\psi_2}}{|1-r^2|}\\
&S_{m_2\psi_2} = \frac{S_{\psi_2\psi_2} - rT_{\psi_1\psi_2}}{|1-r^2|} && T_{m_2\psi_2} = \frac{rS_{\psi_1\psi_2}}{|1-r^2|}\\
&S_{m_1m_2} = \frac{S_{\psi_1\psi_2}}{|1-r^2|} && T_{m_1m_2} = \frac{-r(S_{\psi_1\psi_1} + S_{\psi_2\psi_2}) + (1+r^2)T_{\psi_1\psi_2}}{(1-r^2)^2}\\
&S_{m_1m_1} = \frac{S_{\psi_1\psi_1} + r^2S_{\psi_2\psi_2} - 2rT_{\psi_1\psi_2}}{(1-r^2)^2} && S_{m_2m_2} = \frac{r^2S_{\psi_1\psi_1} + S_{\psi_2\psi_2} - 2rT_{\psi_1\psi_2}}{(1-r^2)^2}
\end{aligned}\right\} \quad (7)$$

Coherence and phase follow from the definitions

$$R_{xy}^2 = \frac{S_{xy}^2 + T_{xy}^2}{S_{xx}S_{yy}}, \qquad \tan\phi_{xy} = \frac{T_{xy}}{S_{xy}}. \quad (8)$$

8. Comparison of observed and computed power

If the excitation spectrum were known, then the wobble spectrum could be computed in accordance with the relations on the last line of equation (7). Note that $S_{m_i m_i}$ becomes infinite at $r = 1$, and so does the total Chandler power $\int S_{m_i m_i}\, df$. This follows because the formulae were developed neglecting dissipation, but very near resonance it is necessary to take the dissipation into account. This is done adding the terms (4) to equation (1) and proceeding as previously. The results are

$$\begin{aligned}
S_{\psi_1\psi_1} &= S_{m_1m_1}+r^2S_{m_2m_2}+2rT_{m_1m_2}-Q^{-1}S_{m_1m_2}+\tfrac{1}{4}Q^{-2}S_{m_2m_2}\\
S_{\psi_2\psi_2} &= S_{m_2m_2}+r^2S_{m_1m_1}+2rT_{m_1m_2}+Q^{-1}S_{m_1m_2}+\tfrac{1}{4}Q^{-2}S_{m_1m_1}\\
S_{\psi_1\psi_2} &= (1-r^2)S_{m_1m_2}+\tfrac{1}{2}Q^{-1}(S_{m_1m_1}-S_{m_2m_2})-\tfrac{1}{4}Q^{-2}S_{m_1m_2}\\
T_{\psi_1\psi_2} &= (1+r^2)T_{m_1m_2}+r(S_{m_1m_1}+S_{m_2m_2})+\tfrac{1}{4}Q^{-2}T_{m_1m_2}.
\end{aligned} \tag{9}$$

To proceed further it is convenient to separate the motion of the pole into positive (west-to-east) and negative motion. At a fixed frequency the relations are

$$\mathbf{m} = m_1+im_2 = \mathbf{m}^+\exp(2\pi if\tau)+\mathbf{m}^-\exp(-2\pi if\tau)$$

and this defines $\mathbf{m}^+(t)$ and $\mathbf{m}^-(t)$. Let S_{m^+} designate the spectrum of $|\mathbf{m}^+|$. It can be shown (Munk & MacDonald 1960, Sections 6.7, 10.1) that for the continuous spectra

$$S_{m^+} = \tfrac{1}{4}(S_{m_1m_1}+S_{m_2m_2}-2T_{m_1m_2})$$

and similarly for S_{ψ^+}. Substituting from equation (9) we find that

$$S_{m^+}(f) = \frac{S_{\psi^+}(f)}{(1-r)^2+\tfrac{1}{4}Q^{-2}}, \qquad r = \frac{f}{f_0}. \tag{10}$$

The expression $[(1-r)^2+\tfrac{1}{4}Q^{-2}]$ is the quadratic impedance of a linear damped resonator.

It is reasonable to assume that $S_{\psi^+}(f)$ is fairly uniform over the narrow range of frequencies contained in the Chandler wobble, so that we can replace it by $S_{\psi^+}(f_0)$. The total power of the Chandler wobble then equals

$$S_{\psi^+}(f_0)\int_{-\infty}^{\infty}\frac{df}{(1-f/f_0)^2+\tfrac{1}{4}Q^{-2}} = 2\pi f_0 Q S_{\psi^+}(f_0).$$

From a fit of the observed spectra, $S_{m^+}(f)$, one finds $f_0 = 0{\cdot}85$ cpy and Q between 10 and 100. The meteorological observations give (in radians^2/cpy)

$$S_{\psi_1\psi_1}(f_0) = 1{\cdot}5\times10^{-16}, \quad S_{\psi_2\psi_2}(f_0) = 0{\cdot}3\times10^{-16}, \quad T_{\psi_1\psi_2}(f_0) = 0{\cdot}06\times10^{-16}$$

for the excitation above the assumed noise level. Thus

$$S_{\psi^+} = 0{\cdot}4\times10^{-16}\ \text{radians}^2/\text{cpy}$$

and the expected Chandler power from meteorologic excitation equals (in radians^2)

Q	10	30	100
Chandler power	$0{\cdot}2\times10^{-14}$	$0{\cdot}6\times10^{-14}$	$2{\cdot}1\times10^{-14}$

as compared to the measured power of $(0''{\cdot}1)^2 = 24\times10^{-14}$ radians^2. Thus the meteorologic excitation fails by one or two orders of magnitude to produce the

observed power of the Chandler wobble. The argument can be put in another form. By choosing a sufficiently high value of Q (order 1 000) it is possible to match computed and observed Chandler power. But in that case the required sharpness of the peak exceeds the measured sharpness; other geophysical considerations also speak against a dissipation as small as is implied by a Q of the order 10^3 (Munk & MacDonald 1960, Section 10.7).

The objection might be raised here that the significance of the excitation spectrum is marginal; this is certainly the case. But then we should expect the computed spectral density to be too high, not too low! In this sense the discrepancy here derived between the computed and observed wobble may be regarded as a lower limit.

9. Coherence and phase

As an independent check we shall consider the coherence and phase relations between the astronomical and meteorological time series. Take, for example, m_1 and ψ_1. All we need to do is to substitute the expressions (7) into the definitions (8). For the case of no friction we find

$$R^2_{m_1\psi_1} = 1 - \frac{r^2 S_{\psi_2\psi_2}(1 - R^2_{\psi_1\psi_2})}{S_{\psi_1\psi_1} + r^2 S_{\psi_2\psi_2} - 2rT_{\psi_1\psi_2}}, \qquad \tan\phi_{m_1\psi_1} = \frac{rS_{\psi_1\psi_2}}{rT_{\psi_1\psi_2} - S_{\psi_1\psi_1}} \tag{11}$$

Thus the predicted coherence diminishes from $R_{m_1\psi_1} = 1$ at $r = 0$ to $R_{m_1\psi_1} = R_{\psi_1\psi_2}$ at $r \gg 1$. The phase varies from $\phi_{m_1\psi_1} = 0°$ at $r = 0$ to $\phi_{m_1\psi_1} = 90° - \phi_{\psi_1\psi_2}$ at large r.

The coherence implied by (11) far exceeds the computed amounts (Figure 5), particularly at small r. The reason for this lies in the present unrealistic model, according to which the loss of coherence is due entirely to imperfect coherence between the components ψ_1 and ψ_2 (if $R_{\psi_1\psi_2} = 1$ then $R_{m_1\psi_1} = 1$ for all r). There are other, much more important, sources for destroying coherence:

(i) Noise in the astronomical observations; but this cannot be a major factor for otherwise coherence would be peaked near f_0 where the noise level is *relatively* low.

(ii) Noise in the meteorologic observations; again this appears not to be the primary source of coherence loss, for otherwise the coherence would be peaked near 1 cpy where the noise level is again relatively small.

(iii) Multiple causes of the Chandler wobble; in fact the comparison of computed and observed spectra demonstrates that this is so and that meteorologic excitation is a rather minor factor.

To pursue (iii) we set

$$\psi_i = \psi_i' + \psi_i'', \qquad R_{\psi'\psi''} = 0$$

where ψ_i' designates the meteorological excitation and ψ_i'' the excitation by some other process uncorrelated with the meteorological excitation. As an example we consider

$$S_{m_1\psi_2} = \frac{S_{\psi_1\psi_2}}{1 - r^2}$$

according to (7). Suppose we define

$$S_{m_1\psi_2'} = \frac{S_{\psi_1'\psi_2'}}{1 - r^2}, \qquad S_{m_1\psi_2''} = \frac{S_{\psi_1''\psi_2''}}{1 - r^2}$$

as the co-spectra of m_1 with ψ_2' and ψ_2'', respectively. Then it follows by the rules in section (7) that $S_{m_1\psi_2} = S_{m_1\psi_2'} + S_{m_1\psi_2''}$ and similarly for other correlations. In particular, $S_{m_im_i} = S_{m_i'm_i'} + S_{m_i''m_i''}$, where $S_{m_i'm_i'}$ is spectrum of m_i due to ψ' and $S_{m_i''m_i''}$ that due to ψ''. The curves in Figure 5 represent functions such as $R_{m_1\psi_1'}$ and $\phi_{m_1\psi_1'}$. If the excitation had been due entirely to meteorological factors $\psi_i = \psi_i'$, $m_i = m_i'$, then equations (11) and other similar expressions apply; if there is an additional, uncorrelated excitation, ψ_i'', then

$$R^2_{m_1\psi_1'} = \frac{S^2_{m_1'\psi_1'} + T^2_{m_1\psi_1'}}{S_{m_1m_1}S_{\psi_1'\psi_1'}} = \frac{S^2_{m_1\psi_1'} + T^2_{m_1\psi_1'}}{S_{m_1'm_1'}S_{\psi_1'\psi_1'}} \frac{S_{m_1'm_1'}}{S_{m_1m_1}}$$

so that $R^2_{m_1\psi_1'}$ is reduced in the ratio $S_{m_1'm_1'}/S_{m_1m_1}$ as compared to the case of atmospheric excitation only. Our previous result that atmospheric excitation accounts for only 1 per cent to 10 per cent of the wobble power then implies a reduction in R by a factor 10 to $\sqrt{10}$. The observed coherence is consistent with such a reduction.*

On the other hand, the computed phase

$$\phi_{m_1\psi_1'} = \arctan(T_{m_1\psi_1'}/S_{m_1\psi_1'})$$

does not depend on whether or not ψ_i'' vanishes, except that its precision rapidly deteriorates with diminishing coherence. According to Goodman (1957) the 95 per cent uncertainty of the computed phase is given by

$$\sin^2\Delta\phi = \frac{1-R^2}{R^2}\frac{6}{\nu}$$

provided $\nu \gg 1$. Even at low resolution the scatter is then very large. Setting $\nu = 100$ (Table 3) and $R = 0{\cdot}1$ gives $\Delta\phi = 50°$. In spite of the unfavourable statistics one is tempted to compare computed and observed phase. The result is disappointing (Figure 7).

The high coherence of m_1 with m_2 at resonance carries no implication concerning the nature of the excitation. Starting with equations (7) it can be shown that

$$R^2_{m_1m_2} = 1 - \frac{(1-r^2)^2 S_{\psi_1\psi_1}S_{\psi_2\psi_2}(1-R^2_{\psi_1\psi_2})}{(S_{\psi_1\psi_1} + r^2S_{\psi_2\psi_2} - 2rT_{\psi_1\psi_2})(r^2S_{\psi_1\psi_1} + S_{\psi_2\psi_2} - 2rT_{\psi_1\psi_2})}$$

and this has a parabolic peak of $R_{m_1m_2} = 1$ at $r = 1$ regardless of whether there are single or multiple causes. The normalized co-spectrum vanishes at resonance, and the quadrature-spectrum is negative. Hence $\phi_{m_1m_2}$ is nearly $-\frac{1}{2}\pi$ at frequencies adjoining f_0, and again this result is independent of excitation.

10. The astronomically inferred excitation spectrum

Figure 8 shows the excitation function as inferred from the astronomical observations in accordance with equations (9). The numerical contributions of the terms in Q^{-1} and Q^{-2} toward the computed excitation is small throughout.

* We have assumed ψ_i'' as completely uncorrelated with ψ_i'. On the other extreme we might contemplate ψ_i'' to be completely coherent with ψ_i'. In that case $R_{m_1\psi_1'}$ would not be reduced, but the computed power would still fail by one to two orders of magnitude to account for the observed power. The present result then implies that the source of the Chandler wobble is not only far more potent than shifts in atmospheric mass but also uncorrelated to it.

The rise of the inferred excitation spectrum at high frequencies is not significant, but merely reflects amplification by a factor $1+r^2$ of the astronomical noise level. At the Chandler frequency the inferred excitation is 3×10^{-14} radians^2/cpy as compared to $1{\cdot}5\times10^{-16}$ from meteorological observations. In Section (8) we compared the computed and observed power of *wobble* and concluded that the meteorological excitation is negligible. The present comparison between the

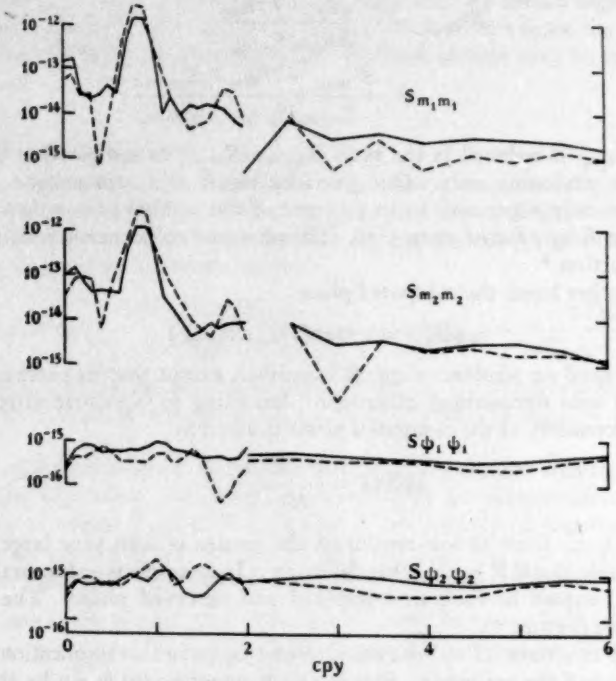


FIG. 6.—The spectra for the periods 1900 to 1920 (dashed) and 1921 to 1950 (solid). Spectra of intermediate resolution were plotted for frequencies below 2 cpy; those of low resolution for frequencies above 2 cpy.

inferred and observed *excitation* leads to the same conclusion, but is subject to far greater statistical uncertainty. The reason is that we are dealing with a small difference between two large numbers. Thus at frequencies of 0·8 and 0·9 cpy, straddling the Chandler frequency, we have

$$S_{m_1m_1}+r^2S_{m_2m_2}-2r(-T_{m_1m_2}) = S_{\psi_1\psi_1}$$

$$r = 0{\cdot}941 : 4503 + r^2(4495) - 2r(4477) = 58$$

$$r = 1{\cdot}059 : 4903 + r^2(4100) - 2r(4064) = 83$$

respectively (all power densities in arbitrary units). In these calculations there are $\nu = 20$ degrees of freedom, and the proportional variance is accordingly $2/\nu = 0{\cdot}1$. Note that the values of $S_{m_1m_1}$ and $S_{m_2m_2}$ are within 0·002 of one another and hence clearly not statistically independent. This makes it a difficult

matter to estimate the significance limits of the inferred excitation spectrum especially near resonance.

For frequencies of less than 2 cpy the coherence between the astronomically inferred ψ_1 and ψ_2 is very high, with ψ_2 leading by close to 90°. But note that $T_{\psi_1\psi_2}$ is a difference between two large numbers; whereas $S_{\psi_1\psi_2}$ is not. Thus the ratio $2_{\psi_1\psi_2}/S_{\psi_1\psi_2}$ is bound to be large in the presence of statistical scatter, especially near resonance, and $\phi_{\psi_1\psi_2} = \arctan(T_{\psi_1\psi_2}/S_{\psi_1\psi_2})$ is likely to be near $\pm 90°$. The preference for $+90°$ in Figure 8 is not understood. It follows from (6) that $S_{\psi_1\psi_1}$, $S_{\psi_2\psi_2}$ and $-T_{\psi_1\psi_2}$ approach identical limits as $r \to 1$, whereas $S_{\psi_1\psi_2} \to 0$. As a consequence $R_{\psi\psi_2} \to 1$ in the presence of errors.

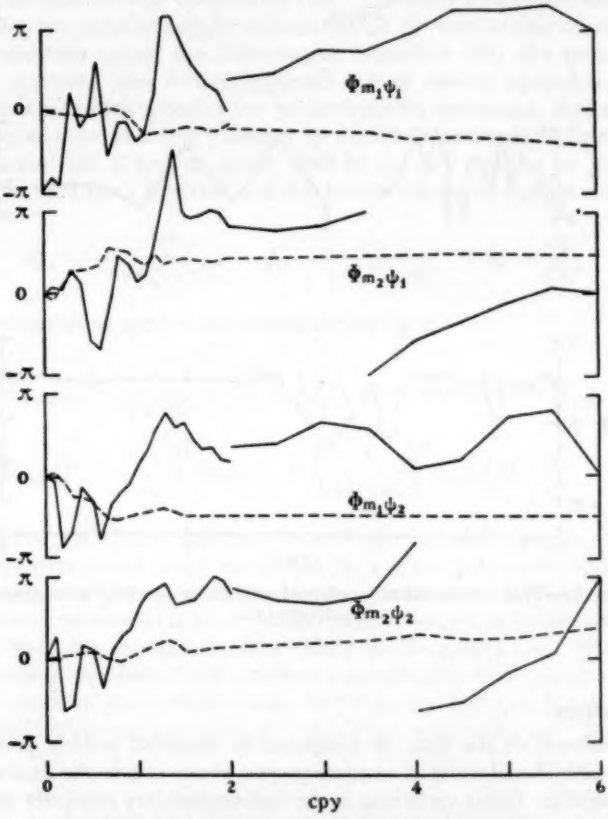


FIG. 7.—Comparison between observed (solid) and computed (dashed) phases. Spectra of intermediate resolution were used for frequencies below 2 cpy; those of low resolution for frequencies above 2 cpy.

A careful analysis of the significance of the inferred excitation would be highly desirable, for it could provide a clue to the nature of the excitation. For example, for a single coherent source we can factor $q(\theta, \lambda; t)$ into $f(\theta, \lambda)g(t)$ in equations (2) so that $\psi_1 = ag(t)$, $\psi_2 = bg(t)$, and hence $R_{\psi_1\psi_2} = 1$. For *two* incoherent sources

$R < 1$. In the case of *randomly* distributed sources the temporal-spacial correlation between $Q(\lambda, t) = \int q(\theta, \lambda; t) \sin^2\theta \cos\theta \, d\theta$ and $Q(\lambda-\delta, t-\tau)$ can be determined (at least in principle) from a knowledge of the excitation spectrum.

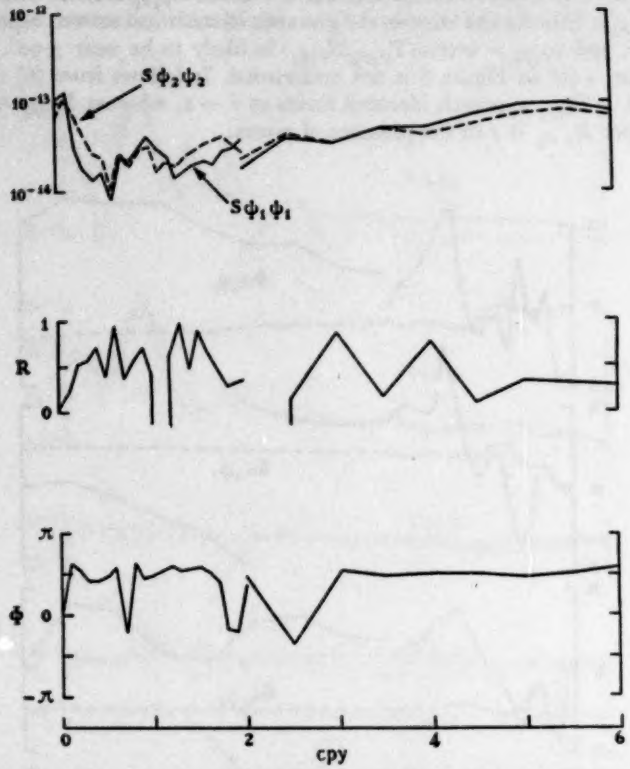


FIG. 8.—The astronomically inferred excitation spectra, according to equations (9).

11. Discussion

The smallness of the ratio of computed to observed wobble power implies that the variable distribution of air mass plays a minor role in the excitation of the Chandler wobble. Other variations in the hydrosphere are relatively unimportant to the *annual* wobble, and can hardly be expected to predominate at 0·85 cpy. Moreover, the low coherence between computed and observed wobble suggests that no other factors in the hydrosphere are of importance. One is tempted (as always) to look to the core.

Equations (1) govern the displacements of the rotation vector, m_i, parallel to the equatorial plane. The third component,

$$m_3 = \frac{\Delta\Omega}{\Omega} = -\frac{\Delta(\text{length of day})}{\text{length of day}}$$

designates the fractional variation in the diurnal rotation. The equation of motion is $m_3 = \dot{\psi}_3$, where ψ_3 is a suitably defined excitation function. Values of ψ_3 can be inferred from astronomical measurements of time, just as ψ_1 and ψ_2 were inferred from measurements of latitude. From comparisons of sidereal and ephemeris time (Munk & MacDonald 1960, Figure 11.10) we estimate

$$\begin{aligned} S_{\psi_3\psi_3} &= 3\times 10^{-15}\,\text{radians}^2/\text{cpy at } f = 0{\cdot}01\,\text{cpy} \\ S_{\psi_3\psi_3} &= 3\times 10^{-17}\,\text{radians}^2/\text{cpy at } f = 0{\cdot}1\,\text{cpy}. \end{aligned} \tag{13}$$

To go further we need to make some geophysical hypothesis concerning ψ_3. The only hypothesis that has survived so far is that it is due to electromagnetic coupling of the mantle with a turbulent core. To check this hypothesis we need to know either the variable angular momentum of the core or the variable flux of angular momentum across the core–mantle boundary (i.e. the torque). The momentum approach and torque approach are of course equivalent, and the choice depends on which quantity can be more readily estimated. Apparently this is the torque, and this is severely limited by the conductivity in the lower mantle. Rochester (1960) has placed an upper limit of 9×10^{24} dyn/cm on the component L_3. The excitation function and the torque are connected by the relations

$$\psi_1 = -\frac{L_2}{\Omega^2(C-A)}, \qquad \psi_2 = \frac{L_1}{\Omega^2(C-A)}, \qquad \dot{\psi}_3 = \frac{L_3}{\Omega C} \tag{14}$$

and the corresponding spectra are related according to

$$\begin{aligned} S_{\psi_1\psi_1}(f) &= \left[\frac{1}{\Omega^2(C-A)}\right]^2 S_{L_1L_1}(f), \quad \text{and similarly for } S_{\psi_2\psi_2} \\ S_{\psi_3\psi_3}(f) &= \left[\frac{1}{2\pi f}\right]^2 S_{\dot{\psi}_3\dot{\psi}_3}(f) = \left[\frac{1}{2\pi f\,\Omega C}\right]^2 S_{L_3L_3}(f). \end{aligned} \tag{15}$$

Taking Rochester's upper limit of 9×10^{24} dyn/cm and assuming a uniform spectrum between 0 and 1 cpy we obtain $S_{L_3L_3}(f) = 0{\cdot}81\times 10^{50}$ (dyn cm)^2/cpy and $S_{\psi_3\psi_3} = 6\times 10^{-19} f^{-2}$ radians^2/cpy, which is twice the estimate in equation (13). This satisfactory agreement is hardly an accident inasmuch as Rochester's estimate is based on the same astronomic observations as the spectra (13). Moreover the atomic frequency standards have provided some detailed information concerning the angular acceleration in recent times. Between July 1955 and January 1958 the length of day has increased at a rate of 0·43 milli-seconds per year. The requisite torque is

$$L_3 = I_m\frac{\Delta\Omega}{\Delta t} = \frac{2\pi I_m}{T^2}\frac{\Delta T}{\Delta t} = 8{\cdot}2\times 10^{24}\,\text{dyn cm}$$

in excellent agreement with the previous estimate. Here I_m is the mantle's moment of inertia and T the length of day.

Suppose we accept $S_{L_3L_3} = 4\times 10^{49}$ (dyn cm)^2/cpy as determined by the length-of-day variations and found not to be inconsistent with electromagnetic theory. From the observed wobble we have estimated $S_{\psi_1\psi_1}$ and $S_{\psi_2\psi_2}$ to be of the order 10^{-15} to 10^{-14} radians^2/cpy. According to (15) the required torque spectra,

$S_{L_1L_1}$ or $S_{L_2L_2}$, then equal 2×10^{53} to 2×10^{54} (dyne cm)^2/cpy. Thus if the core were responsible for both the wobble and the changes in diurnal rotation, the equatorial torque components would have to be several hundred times larger than the axial torque, whereas one expects the equatorial components to be, if anything, less than the axial component. The conclusion is that the core is a negligible source of the Chandler wobble.

Scripps Institution of Oceanography,
University of California,
La Jolla, Calif.,
U.S.A.:
1960 *July* 26.

References

Bartlett, M. S., 1956. *An Introduction to Stochastic Processes* (Cambridge University Press).

Blackman, R. and Tukey, J., 1958. *Bell System Technical Journal*, **37**, 185.

Bullard, E., 1950. *The Observatory*, **70**, 139.

Fellgett, Peter, 1961. (In preparation).

Goodman, N. R., 1957. *New York University, Scientific Paper No. 10*, Engineering Statistics Laboratory.

Hassan, E. S. M., 1960. *Meteorological Monographs* (in press).

Jeffreys, H., 1940. The variation of latitude. *Mon. Not. R. Astr. Soc.*, **100**, 139.

Jeffreys, H., 1959. *The Earth* (Cambridge University Press).

Kempthorne, O., 1957. *An Introduction to Genetic Statistics* (John Wiley & Sons).

Munk, W. H. & MacDonald, G. J. F., 1960. *The Rotation of the Earth* (Cambridge University Press).

Rochester, M. G., 1960. *Phil. Trans. A*, **252**, 531.

Rudnick, P., 1956. *Trans. Amer. Geophys. Un.*, **37**, 137.

Spitaler, R., 1901. *Petermanns Mitteilungen, Erganzungsband*, **29**, 137.

Walker, A. & Young, A., 1957. Further results on the analysis of the observations of the variation of latitude. *Mon. Not. R. Astr. Soc.*, **117**, 119.

Equatorial Upwelling and Sinking in a Zonal Ocean with Lateral Mixing

Koji Hidaka

"The always wind obeying deep."—*Shakespeare, A Comedy of Errors*, I, 1.

Summary

The two-dimensional dynamical equations were solved and an expression for the vertical motion was derived by assuming a balance between the pressure gradient, Coriolis and frictional forces. The frictional forces consist of terms of both vertical and lateral mixing. Isovels of the upwelling and sinking in a meridional section were obtained, assuming the magnitudes of the coefficients of the horizontal and vertical viscosities to be 10^8 and 10^2 c.g.s. respectively. The resulting vertical motion of water seems to be more intense than expected. But the result is quite flexible and reasonable values of vertical motion may be obtained by a proper choice of both the horizontal and vertical viscosity coefficients.

1. Introduction

A few years ago, the present author published a dynamical calculation of the equatorial current system of the Pacific (Hidaka & Nagata 1958) under the assumption that lateral mixing is taking place in addition to the vertical. The calculation was two-dimensional, because the equatorial current system of the Pacific can be approximately regarded as a stationary flow without appreciable variation in EW-direction. Flow patterns in a meridional section were studied. By adopting values 10^2 and 10^8 c.g.s. for the coefficients of vertical and lateral mixing respectively, the authors could explain some of the salient observed natures of the equatorial flow such as an eastward-flowing undercurrent now known as the Cromwell Current after the discoverer.

In that paper, the authors could not succeed in an attempt to calculate the upwelling around the equator, and left this problem to a later study. A solution thereof was recently obtained and the present paper deals with the result of an effort in this direction.

2. Theory

Suppose a meridional section passing through the equator. Take the x-axis positive eastward, the y-axis northward and z-axis vertically downward from the undisturbed level surface. Near the equator the east–west variation in velocity and pressure gradients is generally much smaller than the meridional variation, so

we can neglect the term $\partial^2/\partial x^2$ compared with $\partial^2/\partial y^2$. Then the steady-state equations of motion are as follows:

$$\left.\begin{aligned} A_l\frac{\partial^2 u}{\partial y^2}+A_z\frac{\partial^2 u}{\partial z^2}+2(\omega\rho\sin\phi)v &= \frac{\partial p}{\partial x};\\ A_l\frac{\partial^2 v}{\partial y^2}+A_z\frac{\partial^2 v}{\partial z^2}-2(\omega\rho\sin\phi)u &= \frac{\partial p}{\partial y}, \end{aligned}\right\} \tag{1}$$

where u and v are the x- and y-component of velocity respectively, p is the pressure, ρ the density of sea water, A_l and A_z are the coefficients of horizontal and vertical mixing respectively, ω the angular velocity of the Earth's rotation and ϕ the latitude connected with y by a relation $y = R\phi$. The conditions to be satisfied at the sea surface are

$$z = 0: \quad -A_z\frac{\partial u}{\partial z} = \tau_x; \quad -A_z\frac{\partial v}{\partial z} = \tau_y \tag{2}$$

where τ_x and τ_y are the components of wind stress and generally functions of y or ϕ. Further let us use the bottom conditions:

$$z = h: \quad \frac{\partial u}{\partial z} = \frac{\partial v}{\partial z} = 0 \tag{3}$$

which means that there is no stress at the bottom $z = h$.

If we differentiate both equations of (1) with respect to z and combine the two by a complex expression for the vertical shears,

$$\frac{\partial u}{\partial z}+i\frac{\partial v}{\partial z} = U(y, z) \tag{4}$$

we shall have

$$A_l\frac{\partial^2 U}{\partial y^2}+A_z\frac{\partial^2 U}{\partial z^2}-2i(\omega\rho\sin\phi)U = g\left(\frac{\partial\rho}{\partial x}+i\frac{\partial\rho}{\partial y}\right) \tag{5}$$

with the conditions

$$z = 0: U = -\frac{\tau_x+i\tau_y}{A_z} \tag{6}$$

and

$$z = h: U = 0. \tag{7}$$

In order to solve the equation (5) subject to the conditions (6) and (7), we suppose that there are m stations all spaced at an equal interval Δy on a meridian, and replace in (5) the second derivative of U with respect to y with a finite difference, or

$$\frac{\partial^2 U_k}{\partial y^2} = \frac{U_{k-1}-2U_k+U_{k+1}}{(\Delta y)^2} \tag{8}$$

where U_k and $(\partial^2 U_k)/(\partial y^2)$ represent U and its second derivative at a particular station (kth station) on the meridian under consideration, and U_{k-1} and U_{k+1} refer to the $k-1$th and $k+1$th stations which are located by a spacing Δy to the

north and south of kth station respectively. Then equation (5) becomes the following set of m simultaneous equations:

$$\begin{aligned}
&A_z\frac{d^2U_1}{dz^2}+A_l\frac{U_N-2U_1+U_2}{(\Delta y)^2}-2i\,(\omega\rho\sin\phi_1)U_1=g\left(\frac{\partial\rho}{\partial x}+i\frac{\partial\rho}{\partial y}\right)_1,\\
&A_z\frac{d^2U_2}{dz^2}+A_l\frac{U_1-2U_2+U_3}{(\Delta y)^2}-2i\,(\omega\rho\sin\phi_2)U_2=g\left(\frac{\partial\rho}{\partial x}+i\frac{\partial\rho}{\partial y}\right)_2,\\
&A_z\frac{d^2U_3}{dz^2}+A_l\frac{U_2-2U_3+U_4}{(\Delta y)^2}-2i\,(\omega\rho\sin\phi_3)U_3=g\left(\frac{\partial\rho}{\partial x}+i\frac{\partial\rho}{\partial y}\right)_3,\\
&\qquad\qquad\cdots\cdots\\
&A_z\frac{d^2U_m}{dz^2}+A_l\frac{U_{m-1}-2U_m+U_S}{(\Delta y)^2}-i\,.\,2\omega\rho\sin\phi_mU_m=g\left(\frac{\partial\rho}{\partial x}+i\frac{\partial\rho}{\partial y}\right)_m,
\end{aligned}\tag{9}$$

where U_N and U_S are the values of U at the northern and southern boundaries of the meridional section, and $((\partial\rho/\partial x)+i(\partial\rho/\partial y))_1$, $((\partial\rho/\partial x)+i(\partial\rho/\partial y))_2$, ... , are the values of $(\partial\rho/\partial x)+i(\partial\rho/\partial y)$ at m stations located in between.

Multiply both sides of (9) by a set of m undetermined multipliers l_1, il_2, i^2l_3, ..., $i^{k-1}l_k$, ..., $i^{m-1}l_m$ respectively and add together; then we have

$$\begin{aligned}
&A_z\frac{d^2}{dz^2}(l_1U_1+il_2U_2+i^2l_3U_3+\ldots+i^{m-1}l_mU_m)+\\
&\quad+\frac{A_z}{(\Delta y)^2}\left[\left\{-\left(2+i\,.\,2\omega\rho\frac{(\Delta y)^2}{A_l}\sin\phi_1\right)l_1+il_2\right\}U_1+\right.\\
&\quad+\left\{l_1-\left(2+i\,.\,2\omega\rho\frac{(\Delta y)^2}{A_l}\sin\phi_2\right)l_2+i^2l_3\right\}U_2+\\
&\quad+\left\{il_2-\left(2+i\,.\,2\omega\rho\frac{(\Delta y)^2}{A_l}\sin\phi_3\right)l_3+i^3l_4\right\}U_3+\\
&\quad+\ldots\\
&\quad\left.+\left\{i^{m-2}l_{m-1}-\left(2+i\,.\,2\omega\rho\frac{(\Delta y)^2}{A_l}\sin\phi_m\right)+i^{m-1}l_m\right\}\right]U_m\\
&=-\frac{A_l}{(\Delta y)^2}(l_1U_N+i^{m-1}l_mU_S)+g\left(\frac{\partial\rho}{\partial x}+i\frac{\partial\rho}{\partial y}\right)_1l_1+\\
&\quad+g\left(\frac{\partial\rho}{\partial x}+i\frac{\partial\rho}{\partial y}\right)_2.\,il_2+\ldots+g\left(\frac{\partial\rho}{\partial x}+i\frac{\partial\rho}{\partial y}\right)_m.\,i^{m-1}l_m.
\end{aligned}\tag{10}$$

Here we introduce an unknown ξ which satisfies the following equations:

$$-\left(2+i\,.\,2\omega\rho\frac{(\Delta y)^2}{A_l}\sin\phi_1\right)l_1+il_2=-(2+i\xi)l_1,$$

$$l_1-\left(2+i\,.\,2\omega\rho\frac{(\Delta y)^2}{A_l}\sin\phi_2\right)il_2+i^2l_3=-(2+i\xi)il_2,$$

$$il_2-\left(2+i\,.\,2\omega\rho\frac{(\Delta y)^2}{A_l}\sin\phi_3\right)i^2l_3+i^3l_4 = -(2+i\xi)i^2l_3,$$

. . . . .

$$i^{m-2}l_{m-1}-\left(2+i\,.\,2\omega\rho\frac{(\Delta y)^2}{A_l}\sin\phi_m\right)i^{m-1}l_m = -(2+i\xi)i^{m-1}l_m,$$

or

$$\left.\begin{aligned}
\left(\xi-2\omega\rho\frac{(\Delta y)^2}{A_l}\sin\phi_1\right)l_1+l_2 &= 0,\\
-l_1+\left(\xi-2\omega\rho\frac{(\Delta y)^2}{A_l}\sin\phi_2\right)l_2+l_3 &= 0,\\
-l_2+\left(\xi-2\omega\rho\frac{(\Delta y)^2}{A_l}\sin\phi_3\right)l_3+l_4 &= 0,\\
&\cdots\\
-l_{m-1}+\left(\xi-2\omega\rho\frac{(\Delta y)^2}{A_l}\sin\phi_m\right)l_m &= 0.
\end{aligned}\right\} \tag{11}$$

Eliminating $l_1, l_2, \ldots, l_m$ from these m equations, we shall have an algebraic equation of mth order in ξ. This equation has m roots $\xi_1, \xi_2, \xi_3, \ldots, \xi_m$. For each of these roots, or $\xi_j (j = 1, 2, \ldots, m)$, we have a set of multipliers $l_{1j}, l_{2j}, \ldots l_{mj}$.

Then equation (10) takes the following form:

$$\frac{d^2V_j}{dz^2}-\frac{A_l}{A_z}\,.\,\frac{2+i\xi_j}{(\Delta y)^2}V_j = F_j(z) \tag{12}$$

where

$$\begin{aligned}
V_j &= l_{1j}U_1+il_{2j}U_2+i^2l_3U_3+\ldots+i^{m-1}l_mU_m\\
&= \sum_{k=1}^{m} i^{k-1}l_{kj}U_k
\end{aligned} \tag{13}$$

and

$$F_j(z) = -\frac{A_l}{A_z}\,.\,\frac{l_{1j}U_N+i^{m-1}l_mU_S}{(\Delta y)^2}+\frac{1}{A_z}\sum_{k=1}^{m} i^{m-1}l_{mj}\,.\,g\left(\frac{\partial\rho}{\partial x}+i\frac{\partial\rho}{\partial y}\right)_k. \tag{14}$$

The conditions to be satisfied by V_j can easily be derived as

$$z = 0 : V_j = -\frac{\bar{\tau}_j}{A_z}, \tag{15}$$

where

$$\bar{\tau}_j = \sum_{k=1}^{m} i^{k-1}l_{kj}(\tau_x+\tau_y)_k \tag{16}$$

and

$$z = h : V_j = 0. \tag{17}$$

If we solve the equation (12) subject to the conditions (15) and (17), we have

$$V_j(z) = -\frac{\bar{\tau}_j}{A_z}\cdot\frac{\sinh\lambda_j(h-z)}{\lambda_j\sinh\lambda_j h} - \int_0^h F_j(s)K_j(z,s)\,ds \tag{18}$$

where

$$\left.\begin{aligned} K_j(z,s) &= \frac{\sinh\lambda_j(h-z)\sinh\lambda_j s}{\lambda_j\sinh\lambda_j h} \qquad (s<z)\\ &= \frac{\sinh\lambda_j z\sinh\lambda_j(h-s)}{\lambda_j\sinh\lambda_j h} \qquad (s>z)\end{aligned}\right\} \tag{19}$$

and

$$\lambda_j = +\sqrt{\left(\frac{A_l}{A_z}\cdot\frac{2+i\xi_j}{(\Delta y)^2}\right)}. \tag{20}$$

According to the result of numerical solution of the equations (11), the quantities l_{kj} have largest values on the diagonal ($k=j$) of a matrix $|l_{kj}|$ and decrease rapidly as we go away from the latter, or the larger the value of $|k-j|$, the more l_{kj} becomes smaller.

This means that the term:

$$-\frac{A_l}{A_z}\cdot\frac{l_{1j}U_N+i^{m-1}l_{mj}U_S}{(\Delta y)^2}$$

in (14) has very little influence on the result except at a few stations closest to the northern and southern boundaries of the meridional section. For this reason, we may put $U_N = U_S = 0$ in (14) and use a simpler expression of $F_j(z)$, or

$$F_j(z) = \frac{1}{A_z}\sum_{k=1}^{m} i^{k-1}l_{kj}\cdot g\left(\frac{\partial\rho}{\partial x}+i\frac{\partial\rho}{\partial y}\right)_k \tag{21}$$

provided that we avoid to apply the result to a few stations closest to the northern and southern boundaries.

3. Computation of upwelling

By inversion of the equations (13), we shall have

$$i^{k-1}U_k = \sum_{j=1}^{m}\tilde{l}_{jk}V_j \qquad (k=1,2,\ldots,m) \tag{22}$$

This becomes further

$$U_k = \sum_{j=1}^{m}\frac{\tilde{l}_{jk}}{i^{k-1}}V_j; \tag{23}$$

then we have

$$U_{k+1} = \sum_{j=1}^{m} \frac{l_{j,k+1}}{i^k} V_j.$$

Now the value of $\partial U/\partial y$, at the mid-point between kth and $(k+1)$th station, or $(\partial U_{k+\frac{1}{2}})/\partial y$ is

$$\frac{\partial U_{k+\frac{1}{2}}}{\partial y} = \frac{U_k - U_{k+1}}{\Delta y}$$

$$= \sum_{j=1}^{m} \left(\frac{\bar{l}_{j,k}}{i^{k-1}} - \frac{\bar{l}_{j\,k+1}}{i^k} \right) V_j \tag{24}$$

very accurately.

Since $U = (\partial u/\partial z) + i(\partial v/\partial z)$ by (4) this equation becomes

$$\frac{\partial^2 v_{k+\frac{1}{2}}}{\partial y\, \partial z} + i \frac{\partial^2 U_{k+\frac{1}{2}}}{\partial y\, \partial z} = \sum_{j=1}^{m} \left(\frac{\bar{l}_{jk}}{i^{k-1}} - \frac{\bar{l}_{j,k+1}}{i^k} \right) V_j. \tag{25}$$

Taking the imaginary part only, we have

$$\frac{\partial^2 v_{k+\frac{1}{2}}}{\partial y\, \partial z} = \text{imaginary part of}$$

$$\sum_{j=1}^{m} \left(\frac{\bar{l}_{jk}}{i^{k-1}} - \frac{\bar{l}_{j,k+1}}{i^k} \right) V_j(z) \tag{26}$$

where $V_j(z)$ is given by (18).

From the equation of continuity, we have, for a place not far from the equator,

$$\frac{\partial v}{\partial y} = -\frac{\partial w}{\partial z} \tag{27}$$

where w is the vertical velocity component counted positive downward. Substituting (27) in the left-hand side of (26), we have

$$\frac{d^2 w_{k+\frac{1}{2}}(z)}{dz^2} = \text{imaginary part of}$$

$$\sum_{j=1}^{m} \left(\frac{\bar{l}_{j,k+1}}{i^k} - \frac{\bar{l}_{jk}}{i^{k-1}} \right) \frac{V_j(z)}{\Delta y}. \tag{28}$$

If we integrate this equation subject to the conditions:

$$w = 0 \text{ for } z = 0 \text{ and } z = h, \tag{29}$$

we shall have

$$w_{k+\frac{1}{2}}(z) = \text{ imaginary part of}$$

$$\sum_{j=1}^{m}\left(\frac{\tilde{l}_{j,k+1}}{i^k}-\frac{\tilde{l}_{jk}}{i^{k-1}}\right)\cdot\frac{1}{\Delta y}\cdot\int_0^h V_j(s)G(z,s)\,ds, \tag{30}$$

where

$$\left.\begin{aligned} G(z,s) &= \frac{(h-z)s}{h} && (s<z)\\ &= \frac{z(h-s)}{h} && (s>z)\end{aligned}\right\}. \tag{31}$$

Since $V_j(z)$ is given by (18), we have

$$\begin{aligned}&\int_0^h V_j(s)G(z,s)\,ds\\ &\quad=\int_0^h\left\{-\frac{\bar{\tau}_j}{A_z}\frac{\sinh\lambda_j(h-z)}{\sinh\lambda_j h}-\int_0^h F_j(\sigma)K_j(s,\sigma)\,d\sigma\right\}G(z,s)\,ds\\ &\quad=-\frac{\bar{\tau}_j}{A_z}\int_0^h\frac{\sinh\lambda_j(h-s)}{\sinh\lambda_j h}G(z,s)\,ds-\frac{1}{\lambda_j^2}\int_0^h F_j(s)M_j(z,s)\,ds\end{aligned} \tag{32}$$

where

$$M_j(z,s)=\lambda_j^2\int_0^h G(z,\sigma)K_j(\sigma,s)\,d\sigma. \tag{33}$$

Now that we have

$$\begin{aligned}&\int_0^h\sinh\lambda_j(h-s)G(z,s)\,ds\\ &\quad=\int_0^z\sinh\lambda_j(h-s)\frac{(h-z)\,s}{h}\,ds+\int_z^h\sinh\lambda_j(h-s)\frac{z\,(h-s)}{h}\,ds\\ &\quad=\frac{\sinh\lambda_j h}{\lambda_j^2}\left\{\frac{h-z}{h}-\frac{\sinh\lambda_j(h-z)}{\sinh\lambda_j h}\right\}\end{aligned} \tag{34}$$

and $M_j(z,s)$ becomes for $s<z$ as

$$M_j(z,s)=\lambda_j^2\int_0^s\frac{(h-z)\,\sigma}{h}\cdot\frac{\sinh\lambda_j(h-s)\sinh\lambda_j\sigma}{\lambda_j^2\sinh\lambda_j h}\,d\sigma+$$

then we have

$$U_{k+1} = \sum_{j=1}^{m} \frac{l_{j,k+1}}{i^k} V_j.$$

Now the value of $\partial U/\partial y$, at the mid-point between kth and $(k+1)$th station, or $(\partial U_{k+\frac{1}{2}})/\partial y$ is

$$\frac{\partial U_{k+\frac{1}{2}}}{\partial y} = \frac{U_k - U_{k+1}}{\Delta y}$$

$$= \sum_{j=1}^{m} \left(\frac{\bar{l}_{j,k}}{i^{k-1}} - \frac{\bar{l}_{j\,k+1}}{i^k} \right) V_j \tag{24}$$

very accurately.

Since $U = (\partial u/\partial z) + i(\partial v/\partial z)$ by (4) this equation becomes

$$\frac{\partial^2 v_{k+\frac{1}{2}}}{\partial y\, \partial z} + i \frac{\partial^2 U_{k+\frac{1}{2}}}{\partial y\, \partial z} = \sum_{j=1}^{m} \left(\frac{\bar{l}_{jk}}{i^{k-1}} - \frac{\bar{l}_{j,k+1}}{i^k} \right) V_j. \tag{25}$$

Taking the imaginary part only, we have

$$\frac{\partial^2 v_{k+\frac{1}{2}}}{\partial y\, \partial z} = \text{imaginary part of}$$

$$\sum_{j=1}^{m} \left(\frac{\bar{l}_{jk}}{i^{k-1}} - \frac{\bar{l}_{j,k+1}}{i^k} \right) V_j(z) \tag{26}$$

where $V_j(z)$ is given by (18).

From the equation of continuity, we have, for a place not far from the equator,

$$\frac{\partial v}{\partial y} = -\frac{\partial w}{\partial z} \tag{27}$$

where w is the vertical velocity component counted positive downward. Substituting (27) in the left-hand side of (26), we have

$$\frac{d^2 w_{k+\frac{1}{2}}(z)}{dz^2} = \text{imaginary part of}$$

$$\sum_{j=1}^{m} \left(\frac{\bar{l}_{j,k+1}}{i^k} - \frac{\bar{l}_{jk}}{i^{k-1}} \right) \frac{V_j(z)}{\Delta y}. \tag{28}$$

If we integrate this equation subject to the conditions:

$$w = 0 \text{ for } z = 0 \text{ and } z = h, \tag{29}$$

we shall have

$$w_{k+\frac{1}{2}}(z) = \text{imaginary part of}$$

$$\sum_{j=1}^{m}\left(\frac{\tilde{l}_{j,k+1}}{i^k}-\frac{\tilde{l}_{jk}}{i^{k-1}}\right)\cdot\frac{1}{\Delta y}\cdot\int_0^h V_j(s)G(z,s)\,ds, \tag{30}$$

where

$$\left.\begin{aligned} G(z,s) &= \frac{(h-z)s}{h} && (s<z)\\ &= \frac{z(h-s)}{h} && (s>z)\end{aligned}\right\}. \tag{31}$$

Since $V_j(z)$ is given by (18), we have

$$\begin{aligned}&\int_0^h V_j(s)G(z,s)\,ds\\ &\quad=\int_0^h\left\{-\frac{\bar{\tau}_j}{A_z}\frac{\sinh\lambda_j(h-z)}{\sinh\lambda_j h}-\int_0^h F_j(\sigma)K_j(s,\sigma)\,d\sigma\right\}G(z,s)\,ds\\ &\quad=-\frac{\bar{\tau}_j}{A_z}\int_0^h\frac{\sinh\lambda_j(h-s)}{\sinh\lambda_j h}G(z,s)\,ds-\frac{1}{\lambda_j^2}\int_0^h F_j(s)M_j(z,s)\,ds\end{aligned} \tag{32}$$

where

$$M_j(z,s)=\lambda_j^2\int_0^h G(z,\sigma)K_j(\sigma,s)\,d\sigma. \tag{33}$$

Now that we have

$$\begin{aligned}&\int_0^h\sinh\lambda_j(h-s)G(z,s)\,ds\\ &\quad=\int_0^z\sinh\lambda_j(h-s)\frac{(h-z)\,s}{h}\,ds+\int_z^h\sinh\lambda_j(h-s)\frac{z\,(h-s)}{h}\,ds\\ &\quad=\frac{\sinh\lambda_j h}{\lambda_j^2}\left\{\frac{h-z}{h}-\frac{\sinh\lambda_j(h-z)}{\sinh\lambda_j h}\right\}\end{aligned} \tag{34}$$

and $M_j(z,s)$ becomes for $s<z$ as

$$M_j(z,s)=\lambda_j^2\int_0^s\frac{(h-z)\,\sigma}{h}\cdot\frac{\sinh\lambda_j(h-s)\sinh\lambda_j\sigma}{\lambda_j^2\sinh\lambda_j h}\,d\sigma+$$

$$+\lambda_j^2 \int_s^z \frac{(h-z)\,\sigma}{h} \frac{\sinh\lambda_j s \sinh\lambda_j(h-\sigma)}{\lambda_j^2 \sinh\lambda_j h}\,d\sigma +$$

$$+\lambda_j^2 \int_z^h \frac{z(h-\sigma)}{h} \frac{\sinh\lambda_j s \sinh\lambda_j(h-\sigma)}{\lambda_j^2 \sinh\lambda_j h}\,d\sigma$$

$$= \frac{z(h-s)}{h} - \frac{\sinh\lambda_j(h-z)\sinh\lambda_j s}{\lambda_j \sinh\lambda_j h} \qquad (s < z).$$

Similar expression will be obtained for $M_j(z, s)$ when $s > z$. Thus we have

$$\frac{1}{\Delta y}\int_0^h V_j(s)G(z,s)\,ds$$

$$= -\frac{\bar{\tau}_j}{A_z \Delta y \lambda_j^2}\left(\frac{h-z}{h} - \frac{\sinh\lambda_j(h-z)}{\sinh\lambda_j h}\right)$$

$$-\frac{g}{A_z \Delta y \lambda_j^2}\int_0^h \sum_k i^{k-1} l_{kj}\left(\frac{\partial\rho}{\partial x} + i\frac{\partial\rho}{\partial y}\right)_k . M_j(z,s)\,ds \qquad (35)$$

where $M_j(z, s)$ is a symmetric kernel of the form:

$$M_j(z,s) = \frac{(h-z)s}{h} - \frac{\sinh\lambda_j(h-z)\sinh\lambda_j s}{\lambda_j \sinh\lambda_j h} \qquad (s < z)$$

$$= \frac{z(h-s)}{h} - \frac{\sinh\lambda_j z \sinh\lambda_j(h-s)}{\lambda_j \sinh\lambda_j h} \qquad (s > z).$$

Thus the upwelling velocity $-w_{k+\frac{1}{2}}(z)$ will be, since w is counted positive downwards, given by

$-w_{k+\frac{1}{2}}(z) =$ imaginary part of

$$\sum_{j=1}^{m}\left(\frac{\bar{l}_{j,k+1}}{i^k} - \frac{\bar{l}_{j,k}}{i^{k-1}}\right)\left[\frac{\bar{\tau}_j}{A_z \Delta y \lambda_j^2}\left(\frac{h-z}{h} - \frac{\sinh\lambda_j(h-z)}{\sinh\lambda_j h}\right) + \right.$$

$$\left. + \frac{g}{A_z \Delta y \lambda_j^2} \cdot \int_0^h \left(\frac{\partial\bar{\rho}}{\partial x} + i\frac{\partial\bar{\rho}}{\partial y}\right)_j . M_j(z,s)\,ds\right] \qquad (37)$$

where

$$\bar{\tau}_j = \sum_{k=1}^{m} i^{k-1} l_{kj}(\tau_x + i\tau_y)_k \qquad (38)$$

and

$$g\left(\frac{\partial\bar{\rho}}{\partial x} + i\frac{\partial\bar{\rho}}{\partial y}\right)_j = \sum_{k=1}^{m} i^{k-1} l_{kj} \cdot g\left(\frac{\partial\rho}{\partial x} + i\frac{\partial\rho}{\partial y}\right)_k . \qquad (39)$$

4. A numerical example

The above theory will find a particular application to a number of sections across the equator recently occupied by the oceanographers of the Pacific Oceanic Fishery Investigations, Honolulu, Hawaii. We can use only those sections occupied around the equator, provided the stations are distributed at an equal spacing on the meridian.

For practical computation, we employed the data form the cruise II of the M/V *Hugh M. Smith* of the Pacific Ocean Fishery Investigations, Honolulu, carried out in 1950 January to March, and published in the Special Scientific Report—Fisheries No. 54, Department of the Interior Fish and Wildlife Service. In this cruise two sections were occupied along the meridians of 172° and 158° W, extending from the Hawaiian waters down to 5° S. Only the western section was adopted for computing the equatorial upwelling, the eastern section being used only for computing $\partial\rho/\partial x$ which is actually small compared with $\partial\rho/\partial y$. The stations are distributed at an equal spacing 1° of arc on both of these meridians, one station being located at the equator on each section. Thus we have

$$\begin{aligned}\Delta y &= 111{\cdot}198\,\text{km}\\ &= 1{\cdot}11198\times 10^{7}\,\text{cm}.\end{aligned}$$

For the coefficients of mixing, we adopted $A_l = 10^8$ c.g.s. and $A_z = 10^2$ c.g.s. or $A_l/A_z = 10^6$, as we had in the previous paper (Hidaka & Nagata 1958). The result of numerical computation is given in three diagrams Figures 1, 2 and 3. All these diagrams give the isovels for the vertical motion in a vertical section along the meridian of 172° W. (They are not stream lines.)

5. Discussion

The diagram in Figure 1 was computed from the first term on the right-hand member of the expression (37), and the figures attached to the isovels give the upwelling (negative; shaded areas) and sinking (positive) exclusively due to the action of winds sweeping on the surface of the sea. A strong upwelling is seen exactly at and around the equator, while we have a narrow band of sinking on each side of the former. The upwelling velocity reaches as high as 280×10^{-3} cm/s at a depth of about 200 m, exactly below the equator. The vertical motion rapidly decreases as we go away from the equator to an order of magnitude of 10^{-3} cm/s or smaller. These results suggest that wind action has a remarkable effect on the equatorial upwelling.

Figure 2 gives the upwelling and sinking correlated with mass distribution alone. This was therefore computed from the second term on the right-hand member of the expression (37). It shows a still more intense upwelling, 400×10^{-3} cm/s at a depth of 200 m below the surface exactly at the equator, and in addition to it, sinking of larger downward velocity (500×10^{-3} cm/s) is observed on each side of the former.

Actual upwelling and sinking in the meridional section under consideration must be, however, given by superposing Figures 1 and 2, and the resulting Figure 3 represents the two terms together on the right-hand member of (37). According to this result a strong upwelling amounting to almost 0·5 cm/s is seen at the depth of 200 m exactly below the equator, while the water sinks downward

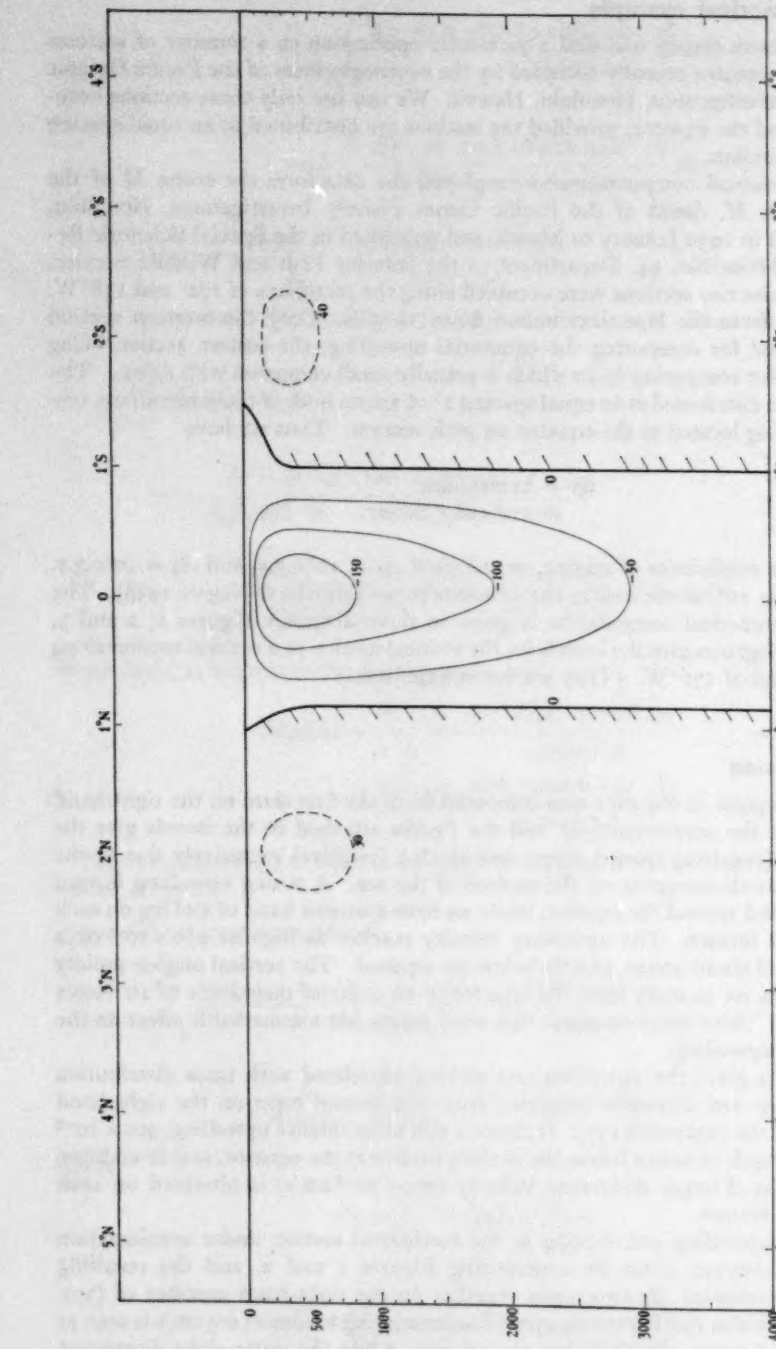


FIG. 1.—Wind-generated upwelling and sinking. (Isovels)
(Unit: 10^{-3} cm/s, vertical scale in metres.)

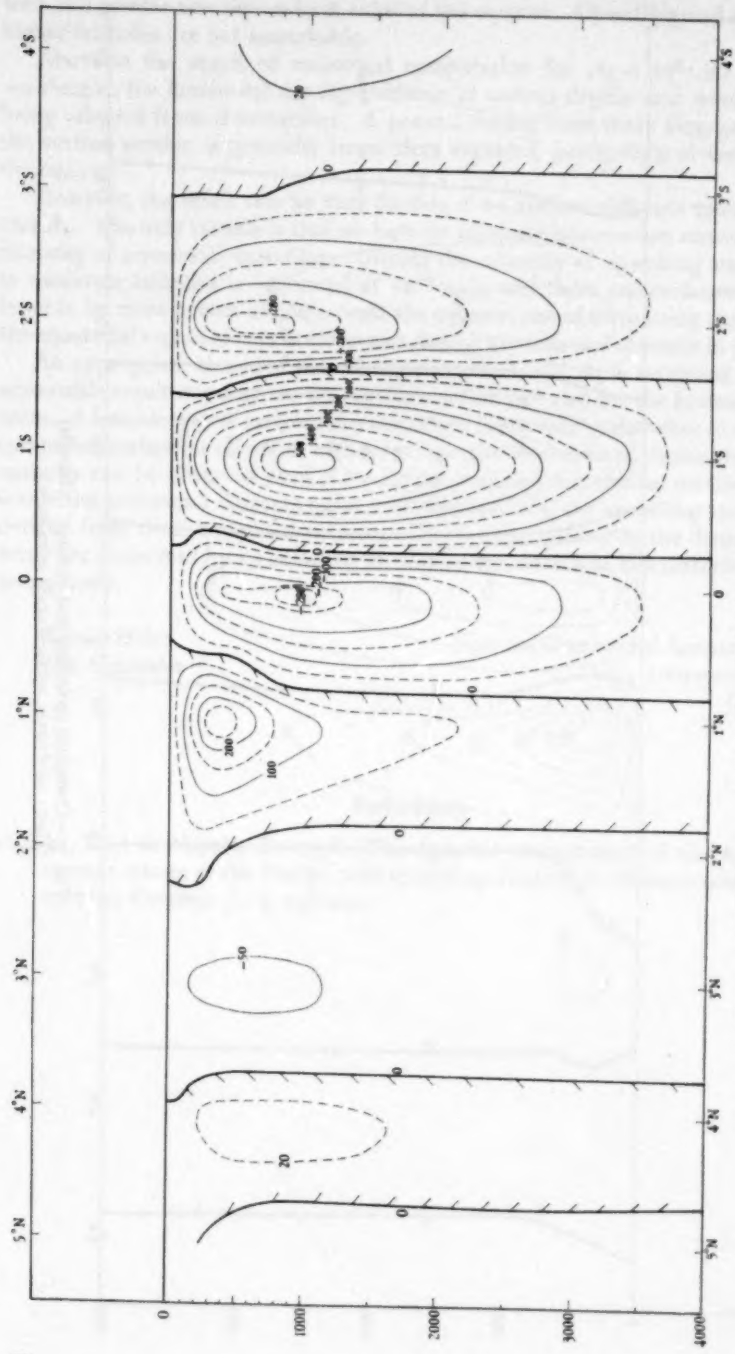


FIG. 2.—Upwelling and sinking correlated with mass distribution. (Isovels)
(Unit: 10^{-3} cm/s, vertical scale in metres.)

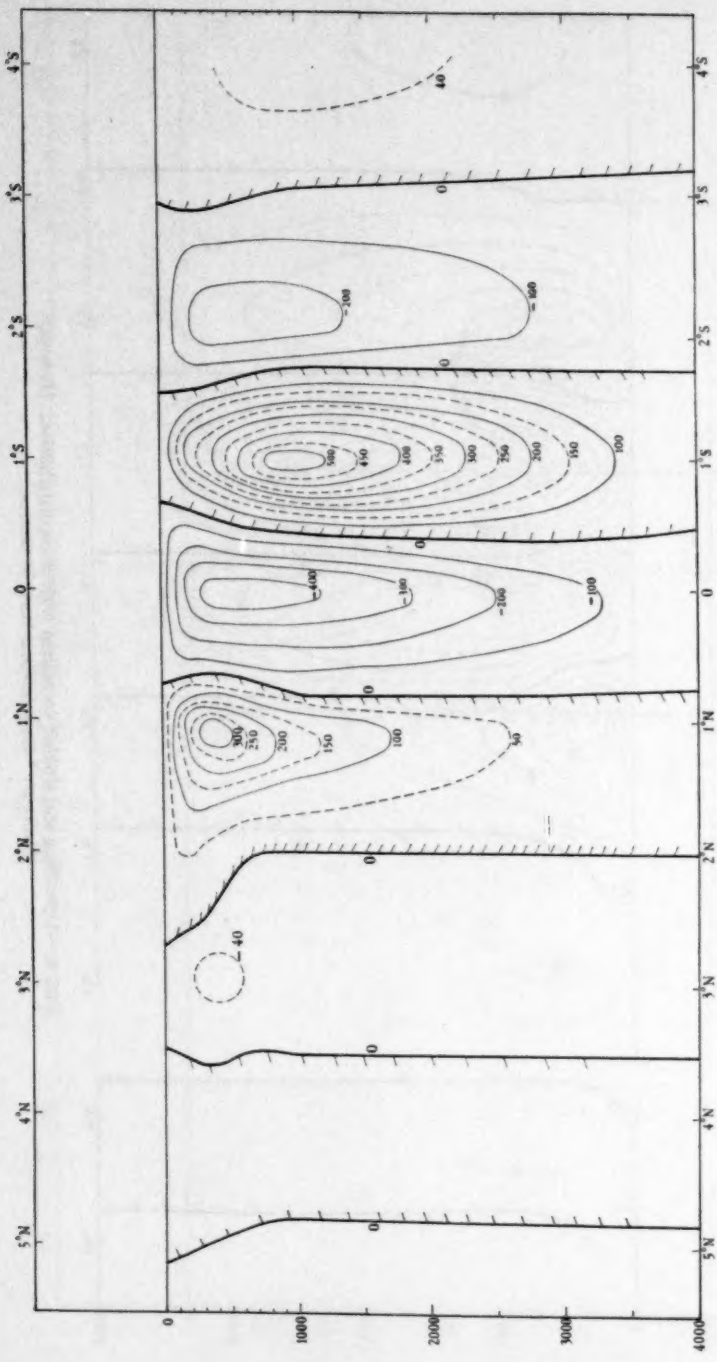


FIG. 3.—Complete upwelling and sinking. (Isovels)
(Unit: 10^{-3} cm/s, vertical scale in metres.)

with still greater velocity on both sides of the equator. Upwelling and sinking at higher latitudes are not remarkable.

Above is the result of numerical computation for $A_l = 10^8$ c.g.s. and $A_z = 10^2$ c.g.s., the horizontal density gradients at various depths and wind stresses being adopted from observations. A general feeling from these diagrams is that the vertical motion is generally larger than expected, particularly at and close to the equator.

However, the result can be very flexible if we assume different values for A_l and A_z. The only trouble is that we have no adequate information concerning the intensity of equatorial upwelling. Usually the intensity of upwelling and sinking in moderate latitudes is estimated at 10^{-3} cm/s, and there are evidences for the latter to be more intense at and around the equator, one of them being the fact that the equatorial region always involves rich fishing grounds and abounds in plankton.

An appropriate choice of the combination of A_l and A_z is supposed to give a reasonable result not only for the vertical motion but also for the horizontal currents. A feasible way of arriving at a consistent result will be therefore to check the computed horizontal velocities with actual current observations, because horizontal currents can be observed even if we cannot measure the vertical motion. If the computed horizontal currents agree with observations, the upwelling and sinking derived from these computed values would be quite reliable as the three components are connected by an equation of continuity. Attack in this direction is now being made.

Woods Hole:
1960 *September* 7.

Now at: *Geophysical Institute,*
Tokyo University,
Japan.

Reference

Hidaka, Koji & Nagata, Y., 1958. The dynamic computation of the equatorial current system of the Pacific, with special application to the equatorial undercurrent. *Geophys. J.*, **1**, 198–207.

VIII. Instruments

J. E. Jackson, The Cambridge pendulum apparatus 375

P. L. Willmore, Some properties of heavily damped electromagnetic seismographs 389

The Cambridge Pendulum Apparatus

J. E. Jackson

"A golden bud set on a leafless stem leads to my result"—*E. Bramah*, Kai Lung Unrolls his Mat.

Summary

The three-pendulum method, developed and used by Lenox-Conyngham, was superseded in 1930, on the arrival of a new vacuum box for swinging two pendulums. With three invar pendulums taken over from the earlier apparatus and three new ones acquired in 1931, this box is the centrepiece of the apparatus which has been in use up to the present day, for geophysical investigations and for establishing gravity reference stations in many parts of the world. This article describes briefly the apparatus and the various modifications of procedure and changes of auxiliary gear that have taken place in the past 30 years. A list of the work done with the apparatus is given.

1. Introduction

For two and a half centuries, geodesists have been studying the Figure of the Earth. Methods of investigation in this field are divisible into two classes, the geometrical and the physical. In the nineteenth century, surveyors with their triangulation and astronomical observations, helped by the utility of their work as a contribution to map-making, got away to a good geometrical start and produced a great many "Figures of the Earth", each one of them closely fitting a small part of the geoid surface.

Geodesy, in the modern sense of the word, stems from Newton's law of gravitation. The variation of gravity with latitude, in a way involved with the Figure of the Earth, and expressed in the famous formula of Clairaut, was the basis for the physical methods of investigation. However, instruments and observations developed slowly on this side. Two reasons for this are obvious enough: the techniques available were inaccurate and tedious, and the results had scarcely any utility value.

The practical method for measuring gravity was the reversible pendulum. Many early observations were made with such instruments. It was soon realized, however, that the direct measurement of gravity acceleration in the field was beset by great difficulties, and that for most geodetic purposes the interesting feature of gravity was its variation from place to place, rather than its actual value anywhere. Out of this came the "invariable pendulum" method. If a pendulum of any kind is taken from place to place and its period of oscillation is measured, the variation of gravity is calculable from the elementary formula $g_2T_2^2 = g_1T_1^2$.

The essential condition is that the pendulum shall not change: more realistically, any change, as that due to a different temperature for instance, is known and can be allowed for. The pendulum may be of quite conventional design, for no linear measurement is needed: indeed, grandfather clocks have been used as instruments for measuring gravity differences.

Special pendulum apparatus for measurement of gravity differences in the field was in use quite early in the nineteenth century. The work was described as "finding the length of the seconds pendulum". Remembering that the range of variation of gravity over the Earth's surface is about 5 gal, we see that an accuracy of a few milligals is required if the physical method is to be comparable in accuracy with the geometrical method for finding the flattening of the Earth's figure. Early differential pendulum work achieved this standard, and the past fifty years have seen a considerable improvement of accuracy, particularly due, of course, to the introduction of radio signals and the use of electronic auxiliary equipment.

Gravity observations have now played their part in the determination of the Figure of the Earth, and this investigation has reached its consummation in Sir Harold Jeffreys's studies and in the International Gravity Formula. (Maybe, the last word has not yet been said on this theme!) However, the observations also detected the anomalous gravity fields connected with structural features of the Earth's crust. Thus, pendulums were swung for geophysical exploration. We recall the discovery of the interesting features of under-sea structure in the East Indies by Vening Meinesz, and the investigation of rift-valley structure by Bullard, as examples.

The co-operation of pendulums with the geological hammer came to an end, however, about twenty years ago when the job of tracking local gravity anomalies was taken over by the much more convenient gravity meters. One might have been tempted to forecast a decline in pendulum swinging, but events have proved to the contrary. Developments during the past twenty years have brought forward the possibility, and the necessity, for a study of the Earth's gravity field as a whole and in some detail. A basic requirement for this study is a world-wide cover of gravity reference stations with values determined to the highest possible accuracy —one milligal or better—for use in calibrating gravity meters and controlling local gravity surveys. The direct measurement of gravity acceleration, to the desired accuracy, is still a slow and costly operation, and pendulum observations are considered to be the most reliable method for measuring large differences of gravity. Establishment of the primary stations is a keen concern of the International Geodetic Association. Several pendulum equipments have contributed to this work, and one such apparatus, the Cambridge Pendulum Apparatus, is the subject of the present article.

Gravity measurement has been actively studied and practised at Cambridge University since 1921, when the Committee for Geodesy and Geodynamics made proposals for the inauguration of a research and teaching branch in these subjects. Sir Gerald Lenox-Conyngham, on his retirement from the Survey of India, brought the "know-how" of his experiences with the three-pendulum apparatus. Much thought and research were devoted to the improvement of method and the design of apparatus, and a new three-pendulum equipment, made by the Cambridge Instrument Company, came into use in 1926. In 1929, a light-alloy box for swinging two pendulums was made by the same firm: this may be regarded as the birth of the apparatus which is still in use.

The basic principle of the method is the determination of the natural periods of two pendulums which swing in antiphase to avoid the sway which would affect the period of a single pendulum: effects of ground motion also are reduced to insignificance by their opposite actions on the two pendulums. For timing the swings and for other necessary observations some auxiliary apparatus is required. Most of this has been designed and constructed in the workshop at Cambridge. It has been modified or rebuilt from time to time, under the expert hands of Mr Leslie Flavill, along with changes of experimental procedure, changes which in some cases led to improvements in accuracy.

The chronological list which follows, makes reference to all the observations done with the Cambridge apparatus, so far as the present author has been able to ascertain. After the list, further details are given about the equipment, methods and results of the work.

The photographs of the vacuum box and pendulum (Figure 1) were supplied by the Cambridge Instrument Company. The other photograph, reproduced by permission of the *Director de Cartografia Nacional*, was taken in Caracas when the apparatus was set up in the Seismological Observatory there (Figure 2).

A plan of the arrangement of the apparatus for observations is seen in Figure 3.

2. Chronological list of the work of the Cambridge two-pendulum apparatus, with some relevant details

1930 Observations by Jolly and Willis, using pendulums I.A and I.C in the new vacuum-box, at Cambridge, Greenwich, Sevenoaks, Wych Cross, Oxford and Bembridge; timing by flash box and chronometer rated against rhythmic time signals: period of I.A apparently increased by about 10^{-6} s during the series.

1931 Observations by Lennox-Conyngham and Willis at Cambridge, Edinburgh and Leith.

Observations by Jolly, Fryer and Cowan in West Scotland and the Isles: 20 stations.

Observations by Willis and Bullard at Greenwich, Southampton and Paris: photographic recording introduced: motor-car accident upset I.C by about 10^{-5} s but it almost recovered in a few days.

New pendulums VI.A, VI.B, VI.C received and adjusted.

Observations by Bullard at Cambridge and Norwich, using VI.A and VI.C.

1932 Observations by Bullard at Cambridge, Bristol, Repton and Mithian, using VI.A and VI.C.

Bullard made extensive tests on the effects of magnetization of pendulums.

Observations by Jolly and others in N. Scotland, Orkney and Shetland using I.A and I.C: 17 stations: period of I.A apparently decreased by about 10^{-6} s.

Tests of new timing method invented by Bullard for comparing the periods of pendulums swung simultaneously at a base station and in the field: pendulum VI.A at Cambridge, I.A and I.C in fields: 4 stations.

Pendulum carrying cases and the vacuum box were lined with thin mu-metal.

1933 Observations by Bullard in N. Wales: 9 stations: pendulums VIA and VI.C at Cambridge base station, I.A and I.C in field: rather large changes of periods of I.A and I.C were noticed.

Observations by Jolly and Bullard at Cambridge, Greenwich and Southampton: large changes in I.A and I.C noticed.

Observations by Lt Bond at 28 stations in S.W. England, with base station at Southampton: large changes in I.A and I.C noticed.

Knife edges of set I pendulums were removed, rust was cleared off and pendulums reassembled.

Bullard took pendulums VI.A, VI.B and VI.C to East Africa for geophysical investigation of the rift valley: pendulums I.A and I.C at Cambridge base, where observations were done by Lenox-Conyngham, Munsey and Flavill: periods of I.A and I.C apparently decreased by about 25×10^{-7} s.

1934 Continuation and completion of the work in East Africa: 56 stations occupied, some of them twice.

Observations by Bullard at Cambridge and Hunstanton; pendulums VI.A and VI.C in field.

Temperature coefficients of VI.A and VI.C re-determined.

1935 Observations at 37 stations in Tanganyika by Horsfield with pendulums VI.A, VI.B and VI.C: pendulums I.A and I.C at Southampton base station.

1936 Observations by Munsey at 6 stations in Sudan; mostly with pendulums VI.A and VI.C: base at Southampton.

Observations by Bullard in Lincolnshire for controlling some torsion-balance work.

1937 Observations by Mace at 13 stations in Cyprus, with pendulums VI.A and VI.C; base station in Department of Geography, Cambridge.

1938 Observations by Kerr-Grant at 34 stations in Australia: base station in Adelaide.

1939 Observations by Browne and Glennie for connection between Cambridge and Dehra Dun; pendulums I.A, I.C, VI.A and VI.C swung at both stations; at Dehra Dun the timing was done with an astronomical clock which was rated by time signals when these could be received.

Observations by Browne and Bullard to connect the "absolute" gravity stations at Teddington and Washington: all six pendulums used at both stations: timing by a new method using laboratory standard frequencies: outbreak of war prevented proposed visit to Potsdam.

1947 Observations by Robertson and Garrick at 19 stations in New Zealand: all six pendulums used: timing done with aid of a crystal oscillator "clock" rated by comparison with a broadcast standard frequency.

1948 Apparatus returned from New Zealand.

Observations by Hales and Gough at 53 stations in southern Africa: apparatus in car accident in October; both sets of pendulums used but only one set taken to field stations.

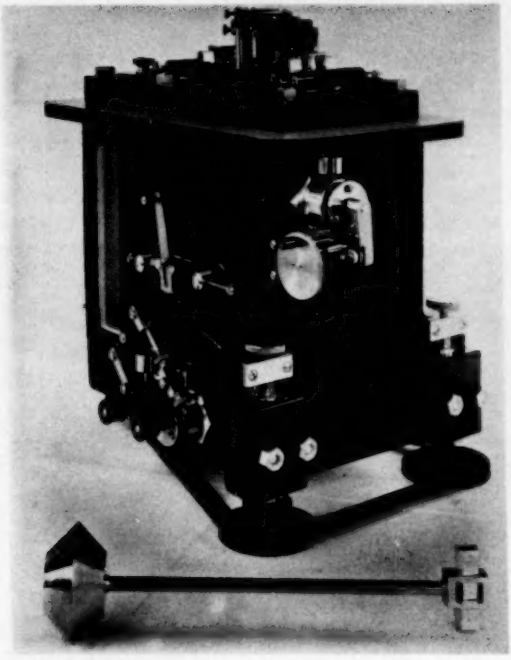
1949 Apparatus returned from South Africa.

Observations by Cook at 4 stations in Eire: pendulums swung in four pairs, I.AB, I.AC, VI.AB and VI.AC.

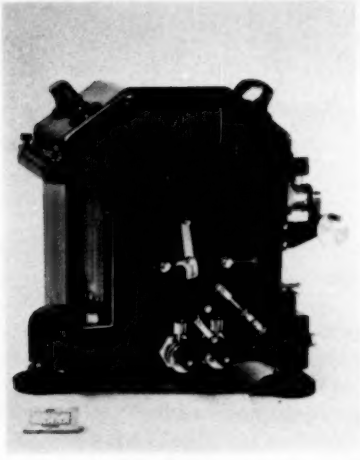
Observations by Browne, Cook, McCarthy and Parasnis at Cambridge, Newcastle, Edinburgh and Aberdeen.

Apparatus taken to Australia.

1950 Observations by McCarthy at 57 stations in Australia.



(a)



(b)

FIG. 1. Views of the case and pendulums.

[facing page 378

FIG. 2. Cambridge pendulum apparatus in use.

1951 Observations by Cook at Teddington and Southampton: anomalous behaviour of pendulum VI.A.

Observations by Cook at Teddington, Sêvres, Brunswick and Bad Harzburg: large changes of periods in set VI, particularly VI.A.

Cook studied theory of magnetic interaction between pendulums and made some experiments.

Apparatus sent to Canada.

1952 Observations by Garland at 10 stations in Mexico and North America: mu-metal lining of box removed and Helmholtz coils used to cancel vertical magnetic field: tests of pendulum magnetization carried out: pendulum VI.A not used.

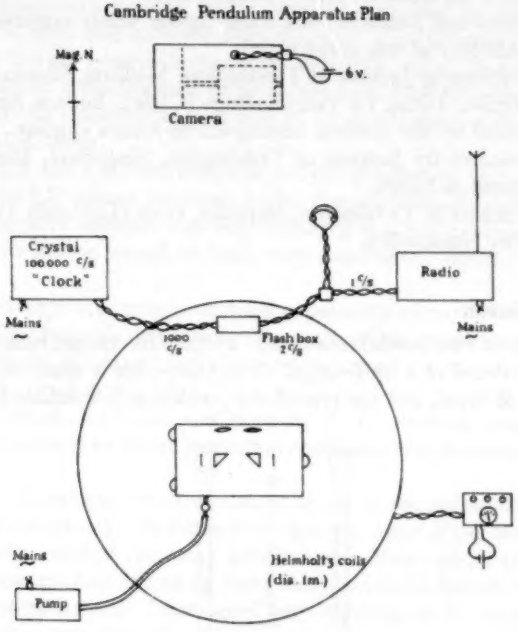


FIG. 3.

1953 Temperature coefficients re-determined by Garland.

Observations by Garland at 10 stations in Canada: pendulum VI.A not used: some erratic results with pendulum I.A.

1953 Garland and Cook made observations for connections between Washing-
1954 ton, Ottawa and Teddington: pendulum VI.A not used.

1954 Observations at Teddington had to be stopped because pendulums failed to keep in step: one agate block found to be loose: all agates removed, refixed and re-lapped.

1955 After the repairs, the pendulums swung satisfactorily but the periods were very anomalous: the dummy pendulum was found to be strongly magnetized: a dummy pendulum of brass was obtained, and all iron components except the knife-edges were replaced by suitable non-ferrous metals. Periods of all six pendulums returned to their normal values.

Crystal clock circuitry rebuilt and a new flash-box made.

Observations were made at Teddington and the apparatus was taken to Oslo. Observations by Jelstrup at Oslo, Bödo, Hammerfest, Copenhagen, Bad Harzburg and Munich: pendulums swung in four pairs, I.AB, I.AC, VI.AB, VI.BC.

1956 Closing observations at Teddington.

1957 Observations by Gough at Mowbray, Pretoria and Johannesburg: pendulums VI.B and VI.C were stained by mercury from a broken manometer, but with no appreciable effect on periods.

1958 Observations by Browne at Teddington, Munich, Rome, Catania, Khartoum, Nairobi, Salisbury and Johannesburg: apparatus travelled by air: most stations occupied on return journey also.

Apparatus and packings were made lighter where possible, and a collapsible Helmholtz coil was constructed.

Observations by Jackson at Teddington, Madison, Mexico City, Panama, Caracas, Quito, Lima, La Paz, Santiago (Chile), Buenos Aires and Rio de Janeiro: travel by air: stations reoccupied on return journey.

1959 Observations by Jackson at Teddington, Singapore, Darwin and Melbourne: travel as before.

1960 Observations at Teddington, Helsinki, Ivalo (Lapland), Hammerfest and Potsdam (by Honkasalo).

3. The pendulums

Figure 4 shows two pendulums drawn to scale in correct relative position for swinging, and a detail of a knife-edge. The knife-edge is made of hard steel, the two wedges are of brass, and the rest of the pendulum is machined from a casting

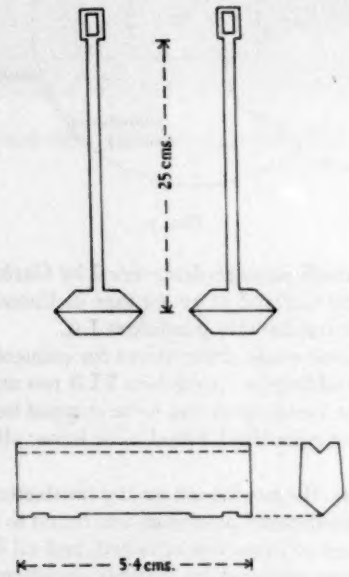


FIG. 4.

of invar metal of composition 64 per cent iron and 36 per cent nickel. The pendulums were designed to have a full period of about 1·01 s, so that they could conveniently be timed by using the succession of coincidences obtainable by comparison with the one-second beats of a chronometer. In fact, the coincidences occur at intervals of about 40 s, depending on the actual value of g at each station.

The set of three pendulums, designated I.A, I.B and I.C, was obtained in 1926, for use in the three-pendulum equipment. Before any field observations could be undertaken, the periods of the pendulums had to be equalized, so that each pair would remain in practically constant phase relationship while swinging together for several hours. Approximate equalization was achieved by the manufacturers, by careful machining. For final adjustment, the pendulums were swung in pairs, and the longer ones were speeded up by carefully regulated grinding of the bottoms of the bobs with fine carborundum powder. This tedious and slow work was done by Sir Gerald Lenox-Conyngham, assisted by Mr J. B. Laws, in 1926.

During the series of field observations in that year, the knife-edge of pendulum I.B fell out, and the equalization job had to be done again, this time with the assistance of Col. Craster.

Pendulums I.A and I.C were involved in a car accident in 1931, with the result that the period of I.C altered by about 10^{-5} s: the observations made immediately after the accident were rejected, but they indicated that the pendulum was rapidly recovering and its period would be back very close to the normal value within a few days.

In the period 1931–33, erratic changes apparently occurred in all the pendulums. Eventually, the knife edges were taken out and traces of rust were found on the contact surfaces. Cleaning and re-assembly cured the trouble.

Another set of pendulums, VI.A, VI.B and VI.C, was obtained in 1931. Final adjustment of periods for these was done by E. C. Bullard, and he also made experimental investigations of the correction coefficients for temperature, buoyancy and amplitude.

Both sets of pendulums were in constant use until the outbreak of war in 1939 stopped this sort of activity. In the period 1932–39, most of the observations were made by Bullard's method requiring simultaneous observations with two sets of apparatus. A new method of timing was being developed before the war, and all six pendulums were used at Teddington and Washington in 1939 for the intercomparison of these two "absolute" stations.

Pendulum work started again in 1947, and the apparatus has been in fairly regular use since then.

Owing to the appearance of some specks of rust, the surfaces of set VI pendulums were lightly gilded: this made no appreciable difference to the periods. The gilt is gradually wearing off.

In the period 1951–53, some rather erratic results gave rise to a suspicion that pendulum VI.A was misbehaving. Subsequent experience indicated that the suspicion was probably unjustified. Possibly, the anomalies were symptoms of the beginning of the mechanical defect which had to be dealt with in 1954.

An accident to a manometer in 1956 caused pendulums VI.B and VI.C to be sprayed with mercury, some of which became amalgamated on to the gilt surfaces. This made no differences to the periods of swing, and the stains are now disappearing.

The graphs in Figure 5 show the half periods of the pendulums as measured at the numerous visits of the apparatus to the gravity station in the National Physical

Laboratory. Dr A. H. Cook has collected some of this information. It seems likely that no serious material changes have occurred in any of the pendulums since about 1933, and that most of the observed changes of period from time to time have come from other sources. It is interesting that the small changes, 1 or 2 parts in 10^6 in 15 yr are less than changes often observed in invar length standards.

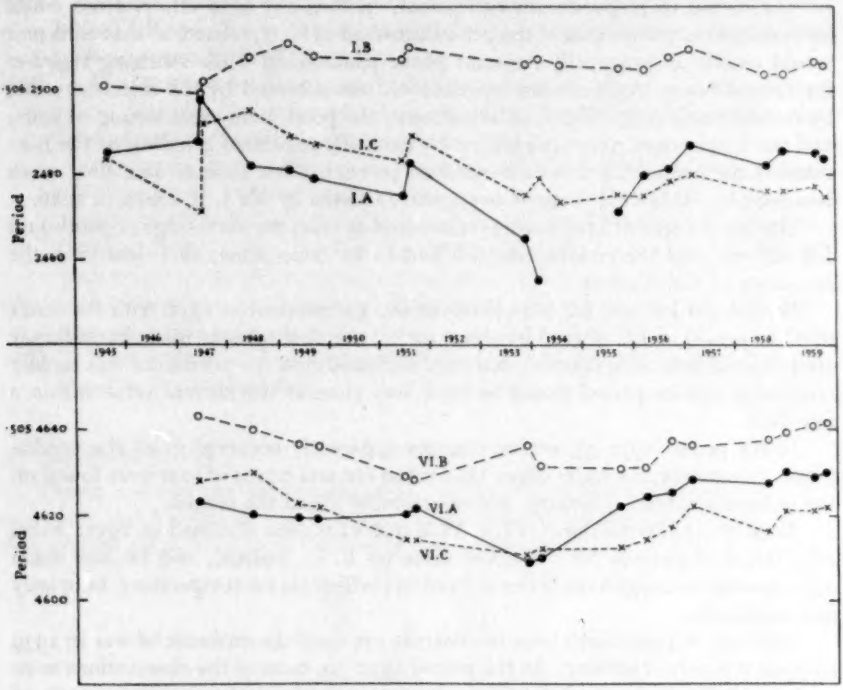


FIG. 5.—Cambridge pendulums. Individual pendulum periods at N.P.L. Values before 1951 are converted from observations at Pendulum House, Cambridge.

4. The vacuum box

This is made of light alloy and is divided into two compartments by an incomplete partition which is high enough to prevent physical interaction between the pendulums through the contained air. There were two taps for pumping purposes, but one has recently been replaced by a plug. The box stands on three short metal legs, two of which are adjustable by screws for levelling and can be firmly clamped when levelling is completed.

Each pendulum swings on two agate blocks and the four blocks are ground and polished so as to be very close to a common plane surface. Coarse levelling of the plane is done with a simple level placed on the flange of the box and the final levelling with a sensitive level in a special fitting which can be placed directly on the agate surfaces. Fixed to the outside of the box there is a bracket with two surface plugs which will take the fine level, so that any dislevellment during a swing can be detected.

A screw and lever arrangement enables the pendulums to be raised from, or lowered on to, the agates. Two axles passing into the bottoms of the two compartments, and connected by levers and linkages, provide for equal deflection of the pendulums before starting a swing, and for stopping the pendulums, when necessary, by bringing small brushes up to touch the under surfaces of the bobs.

Hanging in one compartment is a "dummy" pendulum of about the same mass as a real pendulum: until 1955 it was made of invar. A thermometer is fitted into the stem of the dummy and a manometer is attached: both these instruments can be seen through a rectangular glass window in the side of the vacuum box.

The optical unit with two adjustable prisms stands in a defined position between the two pairs of agates, but is carried separately during transport of the apparatus. When the unit is in position, the pendulums cannot be moved into or out of the box.

The lid of the box fits on to a corresponding horizontal flange. Originally, the vacuum sealing was made by use of heavy grease. This was effective but messy. In 1949 the flanges were cleaned and a rubber O-ring was fitted and found to be quite satisfactory. Observations are normally started at a pressure of about 25 mmHg, and it rarely rises more than 3 mmHg before a pair of swings is completed. There is a glass window in front of the lid so that the light beams can pass in and out.

The trouble in 1954 led to the discovery that one of the agate blocks had become insecure. All the blocks were removed, refixed and ground to a common plane. After further investigations, it was decided to take more drastic action on the effects of magnetization, and all iron in the fittings was replaced by bronze or other suitable metal: this involved the provision of a "dummy dummy pendulum" made of brass and plated. These repairs and modifications were carried out by the Cambridge Instrument Company.

Since the introduction of the Helmholtz coils, it has been necessary to support the vacuum box on three cylindrical brass blocks about 6 cm high, to get the pendulums as nearly as possible central in the coils.

5. Timing methods

For recording the swings of the pendulums, beams of light from two slits are projected into the vacuum box through lenses of one diopter, through the adjustable prisms, then reflected from polished surfaces of the knife-edge blocks on the pendulums and back through the same optical system on slightly different paths. It is not necessary here to describe the flash-box and chronometer system brought over from the earlier apparatus, because in 1931 a photographic method of recording was introduced.

The new method is indicated very schematically in Figure 6. The apparatus is set up and the prisms are adjusted so that when the pendulums are on the agates at rest, the two beams of light come to focus on the lower slit in the camera box. Behind this slit is a cylinder rotated by governed clockwork (gramophone motor) on a threaded axle, at a rate such that its surface moves at about 6½ cm a second. When the pendulums swing, the beams oscillate, normally to about 17 mm above and below the slit, and thus two spots of light are recorded about every 0·505 s on the bromide paper wrapped round the cylinder, as the pendulums pass their rest positions.

Until 1938, the object marked T on Figure 6 was an oscillograph, tuned to respond to a suitable audio frequency. A third beam of light was reflected from the mirror of the oscillograph and brought to focus at a point very close to the camera slit. When the oscillograph was excited by a signal of the appropriate frequency, the light spot oscillated and some light got into the slit. Thus, time signals, or any other timing pulses, could be impressed on the record along with the pendulum swings.

In 1933, Bullard devised his method of comparing two pendulums by recording morse messages. This involved running two sets of pendulums simultaneously, one set at a base station and the other at the field stations. The same morse message was recorded at both places. This is probably the only method by which differences of pendulum periods have been directly measured: all other methods measure pendulum periods.

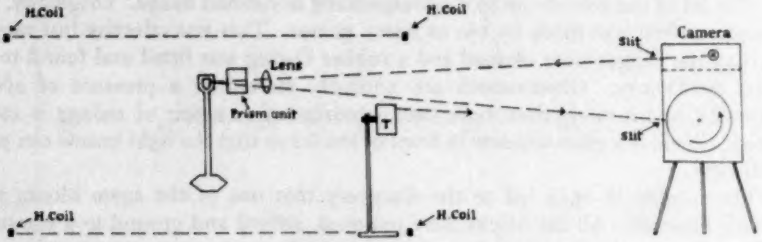


FIG. 6.

The quartz crystal oscillator brought the possibility of a time-measuring instrument of sufficient constancy to be used for measuring pendulum periods in the field to the precision required. A portable frequency standard was successfully used for some submarine gravity observations just before the second world war, and all observations done since the war have been based on this method. The crystal oscillations at 100 kc/s are electronically divided to 1 kc/s and this output drives a synchronous motor. Through gearing the motor drives a rotating disk at 2 revolutions a second: a slit on the disk passes in front of a lamp so that a short flash of light is given out every $\frac{1}{2}$ s. Another disk driven on the same axis at one revolution in 50 s admits extra light for one flash in every hundred. Two counting dials geared in with the mechanism enable each flash to be given a definite number.

This flash-box apparatus takes the place marked T in Figure 6, and it is arranged so that the timing flashes are directed into the camera slit.

Of course a crystal oscillator is regular, but its true time rate must be found by comparison with some kind of time signal. Originally, this was done by loosely coupling the 100 kc/s cycle oscillation with a radio receiver tuned to the B.B.C. Droitwich transmitter, the carrier of which is very accurately regulated to 200 kc/s. The resulting beats could be observed on a milliammeter or other suitable device, and counted with the aid of a watch, so giving a very precise determination of the crystal rate.

Later on the flash-box was modified to take advantage of the continuous transmissions of pulses at 1 s intervals by radio stations like WWV. The new flash-box has an additional disk rotating at 2 cycles a second, with a notch on its edge. A

spring make-break contact is mounted on a loose disk in the same axis, and the rotating disk closes this contact at a particular point in each revolution. The contact is in parallel with headphones on which the WWV pips are being heard, but if the contact is brought round by hand adjustment to a certain position the headphones are short-circuited just as each pulse comes in and no signals are heard. The appropriate position for the contact is easily determined within 1/500 revolution, equivalent to a millisecond. The arrangement is illustrated in Figure 7.

FIG. 7.

The crystal that is normally used in the apparatus runs very closely to its nominal rate, and the contact has to be moved only about 9 ms in 24 hours. Since the timing gear is kept running continuously during a set of observations, at least 40 hours, a very precise determination of its rate is easily made.

It has been suggested that the photographic recording should be abolished in favour of some modern, slick, preferably electronic system. The recording camera runs for about 4 minutes at the beginning, and again 4 minutes at the end of each swing. Each record, once started, is made automatically by the apparatus itself, and contains many more observations than are normally required for the calculations. These features have their values.

6. Magnetism

This is undoubtedly the chief source of anxiety in the use of invar, offsetting the great advantage of its low coefficient of thermal expansion. In 1933, Bullard studied the possible effects of magnetic interaction between the pendulums, measured their magnetic moments, and made experiments to find the actual effects of magnetic interaction and to determine the amount of magnetization that could be tolerated. The vacuum box and the carrying cases for the pendulums were lined with mu-metal $\frac{1}{4}$ mm thick, as a measure to reduce the influence of the Earth's magnetic field.

The magnetic moments of the pendulums were tested at various times during the series of observations made with the apparatus. For the work in North America in 1952, Garland removed the mu-metal and surrounded the vacuum box with a Helmholtz coil carrying the current necessary to cancel the vertical component of the Earth's field. He made frequent observations of the magnetic moments of the pendulums, and found them to remain very steady, at about 30 c.g.s. units except for pendulum VI.C which had about 70 c.g.s. units (pendulum VI.A was not used). Bullard had found that magnetizations of this order, if they remained fairly constant, would have no appreciable effect on differential measurements.

Until 1938, the object marked T on Figure 6 was an oscillograph, tuned to respond to a suitable audio frequency. A third beam of light was reflected from the mirror of the oscillograph and brought to focus at a point very close to the camera slit. When the oscillograph was excited by a signal of the appropriate frequency, the light spot oscillated and some light got into the slit. Thus, time signals, or any other timing pulses, could be impressed on the record along with the pendulum swings.

In 1933, Bullard devised his method of comparing two pendulums by recording morse messages. This involved running two sets of pendulums simultaneously, one set at a base station and the other at the field stations. The same morse message was recorded at both places. This is probably the only method by which differences of pendulum periods have been directly measured: all other methods measure pendulum periods.

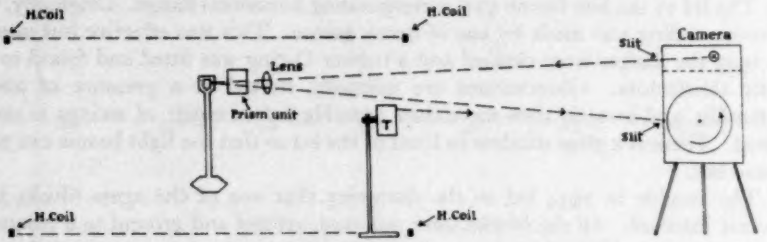


FIG. 6.

The quartz crystal oscillator brought the possibility of a time-measuring instrument of sufficient constancy to be used for measuring pendulum periods in the field to the precision required. A portable frequency standard was successfully used for some submarine gravity observations just before the second world war, and all observations done since the war have been based on this method. The crystal oscillations at 100 kc/s are electronically divided to 1 kc/s and this output drives a synchronous motor. Through gearing the motor drives a rotating disk at 2 revolutions a second: a slit on the disk passes in front of a lamp so that a short flash of light is given out every $\frac{1}{2}$ s. Another disk driven on the same axis at one revolution in 50 s admits extra light for one flash in every hundred. Two counting dials geared in with the mechanism enable each flash to be given a definite number.

This flash-box apparatus takes the place marked T in Figure 6, and it is arranged so that the timing flashes are directed into the camera slit.

Of course a crystal oscillator is regular, but its true time rate must be found by comparison with some kind of time signal. Originally, this was done by loosely coupling the 100 kc/s cycle oscillation with a radio receiver tuned to the B.B.C. Droitwich transmitter, the carrier of which is very accurately regulated to 200 kc/s. The resulting beats could be observed on a milliammeter or other suitable device, and counted with the aid of a watch, so giving a very precise determination of the crystal rate.

Later on the flash-box was modified to take advantage of the continuous transmissions of pulses at 1 s intervals by radio stations like WWV. The new flash-box has an additional disk rotating at 2 cycles a second, with a notch on its edge. A

spring make-break contact is mounted on a loose disk in the same axis, and the rotating disk closes this contact at a particular point in each revolution. The contact is in parallel with headphones on which the WWV pips are being heard, but if the contact is brought round by hand adjustment to a certain position the headphones are short-circuited just as each pulse comes in and no signals are heard. The appropriate position for the contact is easily determined within 1/500 revolution, equivalent to a millisecond. The arrangement is illustrated in Figure 7.

FIG. 7.

The crystal that is normally used in the apparatus runs very closely to its nominal rate, and the contact has to be moved only about 9 ms in 24 hours. Since the timing gear is kept running continuously during a set of observations, at least 40 hours, a very precise determination of its rate is easily made.

It has been suggested that the photographic recording should be abolished in favour of some modern, slick, preferably electronic system. The recording camera runs for about 4 minutes at the beginning, and again 4 minutes at the end of each swing. Each record, once started, is made automatically by the apparatus itself, and contains many more observations than are normally required for the calculations. These features have their values.

6. Magnetism

This is undoubtedly the chief source of anxiety in the use of invar, offsetting the great advantage of its low coefficient of thermal expansion. In 1933, Bullard studied the possible effects of magnetic interaction between the pendulums, measured their magnetic moments, and made experiments to find the actual effects of magnetic interaction and to determine the amount of magnetization that could be tolerated. The vacuum box and the carrying cases for the pendulums were lined with mu-metal $\frac{1}{4}$ mm thick, as a measure to reduce the influence of the Earth's magnetic field.

The magnetic moments of the pendulums were tested at various times during the series of observations made with the apparatus. For the work in North America in 1952, Garland removed the mu-metal and surrounded the vacuum box with a Helmholtz coil carrying the current necessary to cancel the vertical component of the Earth's field. He made frequent observations of the magnetic moments of the pendulums, and found them to remain very steady, at about 30 c.g.s. units except for pendulum VI.C which had about 70 c.g.s. units (pendulum VI.A was not used). Bullard had found that magnetizations of this order, if they remained fairly constant, would have no appreciable effect on differential measurements.

After the repairs to the apparatus in 1955, test observations gave large differences of period for the same pendulum swung in the two positions. This was traced to the dummy pendulum which was found to have a moment of about 450 c.g.s. units. The effect on a pendulum swinging in the same compartment as the dummy, about 8 cm from it, was about 40×10^{-7} s, but the effect on a pendulum in the other compartment was quite small. When all ferrous components of the box, and the dummy pendulum, had been replaced by non-magnetic material in 1955, the pendulum periods returned close to their former values.

A special magnetometer has now been constructed, with which the pendulums are tested immediately before swings are made, and any pendulums that have acquired intolerable magnetization are degaussed in a simple solenoid. Usually it is found that one or two of the pendulums require treatment at each station, not always the same pendulums. As an exception, all pendulums acquired considerable magnetization during their flight from Singapore to London in 1959.

In trying to find out when the magnetization of the old dummy occurred, one may look at the changes of periods of the pendulums at one place from time to time, at Teddington, for example, and one may study the differences of the periods of each pair swung together. There is not room here to quote all the figures, but there seems to have been a small discontinuity in the behaviour of Set I pendulums in 1948 and there were a few curious results with Set VI in 1951. (For the present, we disregard the large effects of the purely mechanical trouble in 1954.) The periods of I.A and VI.A at Teddington have been practically the same, since 1955, as they were in 1951, apart from the anomalous results with Set VI mentioned above and a small but perhaps significant change in both I.A and VI.A in February 1954, when the apparatus returned from Canada; these are the two pendulums which were swung in the compartment with the dummy. The evidence seems to indicate that the strong magnetization of the dummy pendulum was acquired in 1954 or 1955, otherwise its presence would have been evident in earlier observations.

7. Accuracy

When the apparatus was constructed, an accuracy of one milligal was hoped for, and seems to have been achieved. Very soon, however, the figure improved to 0·5 or 0·3 mgal. It is to be expected that increase of precision will come from several sources, such as the accumulated experience of observers, the running-in of the pendulums, the greater care taken over magnetization effects, and the improved facilities for time measurement. Observation in recent years has given, on internal evidence, the standard deviation of about 0·1 mgal for the calculated difference of gravity between successive stations. The procedure now adopted, of traversing each series of stations in both directions, and swinging four different pairs of pendulums, has the valuable feature of providing eight measures of each difference of gravity.

Department of Geodesy and Geophysics,
Cambridge University:
1960 *September* 15

References
(in chronological order)

Bullard, E. C., 1933. The observation of gravity by means of invariable pendulums. *Proc. Roy. Soc. A*, **141**, 233.

Bullard, E. C., 1933. The effect of a magnetic field on relative gravity determinations with Invar pendulums. *Proc. Camb. Phil. Soc.*, **29**, 288.

Bullard, E. C. and Jolly, H. L. P., 1936. Gravity measurements in Great Britain. *Mon. Not. R. Astr. Soc. Geophys. Suppl.*, **3**, No. 9, 443.

Bullard, E. C., 1936. Gravity measurements in East Africa. *Phil. Trans. Roy. Soc. A*, **235**, 445.

Horsfield, W. & Bullard, E. C., 1937. Gravity measurements in Tanganyika Territory, carried out by the Survey Division, Department of Lands and Mines. *Mon. Not. R. Astr. Soc. Geophys. Suppl.*, **4**, 94.

Munsey, D. F. & Bullard, E. C., 1937. Gravity measurements in the Anglo-Egyptian Sudan. *Mon. Not. R. Astr. Soc. Geophys. Suppl.*, **4**, 114.

Mace, C., 1939. Gravity measurements in Cyprus. *Mon. Not. R. Astr. Soc. Geophys. Suppl.*, **4**, 473.

Glennie, E. A., 1939. *Survey of India Geodetic Report for* 1939: Chapter II.

Browne, B. C. & Bullard, E. C., 1940. Comparison of the acceleration due to gravity at the National Laboratory, Teddington and the Bureau of Standards, Washington, D.C. *Proc. Roy. Soc., A*, **175**, 110.

Hales, A. L. & Gough, D. I., 1950. Measurements of Gravity in Southern Africa. (Government Printer, Pretoria.)

Cook, A. H., 1950. Measurements of gravity in Ireland: pendulum observations at Dublin, Sligo, Galway and Cork. *Dublin Institute for Advanced Studies, Geophys. Mem.*, No. 2, Part I.

Browne, B. C., Cook, A. H., McCarthy, E. J. & Parasnis, D. S., 1950. Gravity measurements at York, Newcastle-upon-Tyne, Edinburgh and Aberdeen. *Mon. Not. R. Astr. Soc. Geophys. Suppl.*, **6**, 91.

Cook, A. H., 1952. A measurement of the difference of gravity between the National Physical Laboratory, Teddington and the Ordnance Survey Office, Southampton. *Mon. Not. R. Astr. Soc. Geophys. Suppl.*, **6**, 320.

Cook, A. H., 1952. Comparison of the acceleration due to gravity at the National Physical Laboratory, Teddington, the Bureau International des Poids et Mesures, Sèvres, the Physikalisch-Technische Bundesanstalt, Brunswick, and the Geodetic Institute, Potsdam. *Proc. Roy. Soc., A*, **213**, 408.

Garland, G. D., 1953. Gravity measurements in N. America with the Cambridge pendulum apparatus. *Proc. Roy. Soc., A*, **219**, 215.

Cook, A. H., 1954. Magnetic perturbations of Invar pendulums used for relative gravity measurements. *Mon. Not. R. Astr. Soc. Geophys. Suppl.*, **7**, 1.

Garland, G. D., 1955. Gravity measurements in North America with the Cambridge pendulum apparatus II. *Proc. Roy. Soc., A*, **233**, 203.

Garland, G. D. & Cook, A. H., 1955. A determination of the differences in gravity between the National Physical Laboratory, Teddington, the Dominion Observatory, Ottawa, and the National Bureau of Standards, Washington. *Proc. Roy. Soc., A*, **229**, 445.

Jelstrup, G., 1957. Observations on the gravimetric-calibration base Hammerfest–Munich with the Cambridge pendulum apparatus. *Norges geografiske oppmåling, Geodetic Publication*, No. 7, Oslo, 1957.

Cook, A. H., 1957. Comparisons between relative gravity measurements with the Cambridge pendulum apparatus and Worden gravity meters in North America, South Africa, Australia, New Zealand, and Great Britain. *Mon. Not. R. Astr. Soc. Geophys. Suppl.*, **7**, 228.

Gilbert, R. L. G., 1957. Some comments on the results obtained with the Cambridge pendulum apparatus in North America. *Trans. Amer. Geophys. Un.*, **38**, 457.

Gough, D. I., 1958. A new determination of the differences in gravity between the N.P.L. Teddington and the South African Fundamental and Secondary Stations. *Geophys. J.*, **1**, 298.

Jackson, J. E., 1959. Observations with the Cambridge pendulum apparatus in North, Central and South America in 1958. *Geophys. J.*, **2**, 337.

Jackson, J. E., 1960. Pendulum observations at Teddington, Singapore, Darwin and Melbourne in 1959. *Geophys. J.*, **3**, 78.

Robertson, H. W. & Garrick, R. A., 1960. Gravity measurements in New Zealand with the Cambridge pendulum apparatus (in press). *N.Z. J. Geol. & Geophys.* (This will describe the 1947 observations).

Some properties of Heavily Damped Electromagnetic Seismographs

P. L. Willmore

"My name means the shape I am, and a good handsome shape it is too. With a name like yours you might be any shape almost." *Lewis Carroll*, Alice Through the Looking-Glass

Summary

The classical theory of electromagnetic seismographs is discussed with special reference to the properties of combinations of seismometer and galvanometer in which the damping of either or both components in the system is very much greater than critical. The paper includes a graphical method of determining approximate response curves for such seismographs.

The treatment of passive systems is extended to cover cases in which the galvanometer light spot is made to energize a photoelectric amplifier, and in which the amplifier output is fed back into the mechanical system. It is shown that feedback into the galvanometer changes the apparent period or damping constant, but that such changes can only lead to types of seismograph which could, in principle, be constructed without the use of amplifiers. Feedback into the seismometer does, however, produce new effects. In particular it is shown that feedback proportional to the velocity of the galvanometer spot is equivalent to a change in the galvanometer reaction, such that the reaction constant λ may be given any desired positive or negative value. Large positive values of λ, in conjunction with large damping coefficients, lead to extremely wide-band characteristics. The argument is illustrated by experimental data on a seismograph constructed in Ottawa which yields an almost flat response to ground velocity from 0·5 s to 800 s period, and it is concluded that this pass-band could, if necessary, be extended by at least a decade at each end.

1. Introduction

An electromagnetic seismograph consists of two mechanical oscillators (the seismometer mass and the galvanometer coil) coupled together electrically. If either system is clamped, and the other is deflected from its mean position, the deflected mass will return to equilibrium in a series of damped oscillations if the damping is less than critical, or will return exponentially if the damping is more than critical. In the over-damped case, the return towards the original zero may be quite slow and it will not be possible to determine the free period of oscillation of the displaced system simply by observing the decay characteristic.

These facts have led to some confusion amongst seismologists, for it is widely believed that the inclusion of over-damped components in a seismograph will necessarily lead to a "sluggish" response to the Earth movement, or that over-damped systems have "indefinite" frequency characteristics. Such beliefs are quite erroneous and we shall show that the use of heavily damped components is fully consistent with sharply defined band-pass characteristics, which may permit a good response to Earth movements having frequencies much higher or much lower than the natural frequencies of the seismometer and galvanometer suspensions. To do this we shall first summarize the ordinary theory of the electromagnetic seismograph, we shall then develop an approximate graphical treatment which enables the salient features of the response curve to be predicted intuitively, and we shall finally consider some of the variations which can be introduced by the use of electronic amplifiers and feedback circuits.

2. Theory of the electromagnetic seismograph

The theory of the electromagnetic seismograph was developed in stages by Galitzin (1914), by Wenner (1929) and by Grenet & Coulomb (1935). Following on the lines indicated by these authors, we imagine the seismometer and galvanometer to be connected through a general 4-terminal network, which we replace by a "T" network of 3 resistors as shown in Figure 1. The combined resistance of

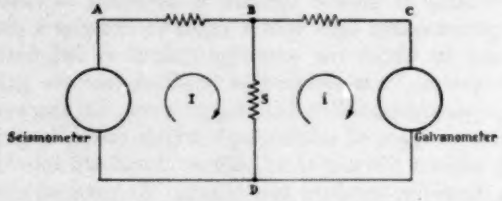


FIG. 1—General 3-terminal resistance network, coupling a seismometer to a galvanometer.

the arm AB and the seismometer coil is called R, the resistance of BC and the galvanometer coil is called r, and the shunt resistance is called S. Currents I and i flow through the seismometer and galvanometer coils respectively.

The equation of motion of the seismometer boom is:

$$K\frac{d^2\phi}{dt^2}+D\frac{d\phi}{dt}+U\phi = LM\frac{d^2x}{dt^2} = GI \qquad (1)$$

where ϕ is the boom deflection, K, D, and U are the moment of inertia, the damping constant and the spring stiffness, M is the mass of the boom, L is the distance from the centre of gravity of the boom to the hinge, x is the displacement of the ground and G is the transducer constant.

The equation of motion of the galvanometer coil is:

$$k\frac{d^2\theta}{dt^2}+d\frac{d\theta}{dt}+u\theta = gi \qquad (2)$$

where θ is the angle of rotation, and k, d, u and g correspond to the quantities K, D, U and G in the seismometer.

The e.m.f.s in the meshes ABD and BCD of Figure 1 will be $G(d\phi/dt)$ and $-g(d\theta/dt)$ respectively, where the negative sign of the galvanometer e.m.f. indicates that it tends to oppose the flow of current. Solving for the currents I and i, we find

$$I = \frac{G(r+s)}{Q^2}\frac{d\phi}{dt}+\frac{gS}{Q^2}\frac{d\theta}{dt}. \tag{3}$$

and

$$i = \frac{GS}{Q^2}\frac{d\phi}{dt}-\frac{g(R+S)}{Q^2}\frac{d\theta}{dt}. \tag{4}$$

where

$$Q^2 = Rr+RS+rS.$$

Substituting the expressions for I and i into equations (1) and (2), and putting

$$\Delta = D+G^2\frac{(r+S)}{Q^2}, \qquad \delta = d+g^2\frac{(R+S)}{Q^2}$$

$$\omega_0^2 = u/k, \qquad \Omega_0^2 = U/K, \quad \alpha = \delta/2k\omega_0, \quad \beta = \Delta/2K\Omega_0$$

$$LM/K = 1/l, \quad \sigma_s = gGS/\Delta Q^2, \quad \sigma_g = gGS/\delta Q^2,$$

the equations become

$$\frac{d^2\phi}{dt^2}+2\beta\Omega_0\frac{d\phi}{dt}+\Omega_0^2\phi = \frac{1}{l}\frac{d^2x}{dt^2}+2\beta\Omega_0\sigma_s\frac{d\theta}{dt} \tag{5}$$

and

$$\frac{d^2\theta}{dt^2}+2\alpha\omega_0\frac{d\theta}{dt}+\omega_0^2\theta = 2\alpha\omega_0\sigma_g\frac{d\phi}{dt}. \tag{6}$$

Eliminating ϕ between equations (5) and (6), we find

$$\begin{aligned}\frac{d^4\theta}{dt^4}+\frac{d^3\theta}{dt^3}(2\alpha\omega_0+2\beta\Omega_0)+\frac{d^2\theta}{dt^2}(\omega_0^2+4\alpha\beta\omega_0\Omega_0+\Omega_0^2-4\alpha\beta\omega_0\Omega_0\sigma_g\sigma_s)+\\ +\frac{d\theta}{dt}(2\beta\Omega_0\omega_0^2+2\alpha\omega_0\Omega_0^2)+\theta\omega_0^2\Omega_0^2 = \frac{2\alpha\omega_0\sigma_g}{l}\frac{d^3x}{dt^3}.\end{aligned} \tag{7}$$

To obtain the periodic solution of equation (7), we set x equal to the real part of

$$X = x_0\exp(j\omega t) \tag{8}$$

and θ equal to the real part of

$$\Theta = \theta_0\exp j(\omega t+\gamma). \tag{9}$$

Equation (7) is regarded as the real part of a complex equation in Θ and X, the solution of which is

$$\begin{aligned}\Theta\left(\frac{dX}{dt}\right)^{-1} = \frac{-2\alpha\sigma_g}{l\Omega_0}\left\{\left[\frac{\omega^2}{\omega_0\Omega_0}+\frac{\omega_0\Omega_0}{\omega^2}-\frac{\omega_0}{\Omega_0}-\frac{\Omega_0}{\omega_0}-4\alpha\beta(1-\sigma_g\sigma_s)\right]-\right.\\ \left.-j\left[2\alpha\left(\frac{\omega}{\Omega_0}-\frac{\Omega_0}{\omega}\right)+2\beta\left(\frac{\omega}{\omega_0}-\frac{\omega_0}{\omega}\right)\right]\right\}^{-1}.\end{aligned} \tag{10}$$

Having solved the equation of motion, we ask ourselves whether the seismometer and galvanometer constants are determined when the equation of motion is given. Coulomb and Grenet solved this problem by a further transformation of equation (7). Setting

$$\omega_0\Omega_0 = \omega'^2, \quad \Omega_0/\omega_0 = \rho^2 \quad \text{and} \quad \omega' t = \tau$$

they obtain

$$\frac{d^4\theta}{d\tau^4} + A\frac{d^3\theta}{d\tau^3} + B\frac{d^2\theta}{d\tau^2} + C\frac{d\theta}{d\tau} + \theta = F\frac{d^3x}{d\tau^3} \tag{11}$$

where

$$\begin{aligned} A &= 2(\beta\rho + \alpha/\rho) \\ B &= \rho^2 + \rho^{-2} + 4\lambda\alpha\beta \\ C &= 2(\beta/\rho + \alpha\rho) \\ \lambda &= 1 - \sigma_s\sigma_g \end{aligned} \tag{12}$$

and F is a constant which determines the absolute magnification. The values of σ_s and σ_g are such that $1 > \lambda > 0$, with $\lambda = 1$ corresponding to the case in which the galvanometer reaction is negligible.

Eliminating α and β in equations (12) leads to the equation

$$(a^2-4)(a-B) + \lambda AC(a - A/C - C/A) = 0 \tag{13}$$

where $a = (\rho^2 + \rho^{-2})$. This cubic equation in a will have either one or three real roots for each value of λ, and each real root will lead to a pair of seismographs, one of which may be converted to the other by interchanging the constants of the seismometer and galvanometer.

In mathematical terms, the problem is now solved, but unfortunately few practical seismologists have much facility with mathematics and few mathematicians have concerned themselves with the little problems of practical seismology. In consequence, many potentially useful combinations have been left untested, and the heavily damped members of this neglected family are some of the most interesting.

In the next section, we shall develop an approximate treatment for overdamped seismographs for which $\lambda = 1$, i.e. the seismographs for which the galvanometer reaction is negligible. Such combinations necessarily have loose coupling and are thus not the most efficient ones which can be made, but the discussion does provide considerable insight into the properties of all seismographs.

In order to determine the circuit characteristics required to maintain an acceptable value of λ, we write

$$\lambda = 1 - \frac{(\Delta - D)(\delta - d)}{\delta\Delta} \cdot \frac{S^2}{(R+S)(r+S)}. \tag{14}$$

Usually $R \geqslant r$, so that $S^2/(R+S)(r+S)$ is less than the square of the d.c. transfer factor of the coupling network, so that attentuation of $1/2$ will ensure that λ exceeds $0{\cdot}75$. In most cases, the difference between $\lambda = 0{\cdot}75$ and $\lambda = 1$ would not produce any significant difference in the ability of a seismograph to record a particular component of Earth movement.

3. Approximate graphical representation of the equations

To develop our approximate approach, set $\lambda = 1$ in equation (10), and then factorize as follows:

$$\Theta = \frac{2\alpha\sigma_g(dX/dt)}{l\Omega_0[j\omega/\omega_0 + 2\alpha - j\omega_0/\omega][j\omega/\Omega_0 + 2\beta - j\Omega_0/\omega]}$$

$$= gGSLM\frac{dX}{dt}\Big/ Q^2 Z_m Z_g \qquad (15)$$

where Z_m and Z_g are the mechanical impedances of the suspended systems of the seismometer and galvanometer respectively. This result could have been written down from first principles, for $LM(d^2X)/(dt^2)$ is the moment of the force which, acting on the boom, is equivalent to the effect of the Earth's acceleration. Dividing by Z_m yields the angular velocity of the boom, multiplying by $(GS)/Q^2$ gives the current through the galvanometer, multiplication by g gives the couple on the galvanometer coil, and division by Z_g gives the angular velocity of the coil. Thus we would have derived the relation between $d\Theta/dt$ and d^2X/dt^2, which is the time differential of equation (15).

Let us now split the overall sensitivity into its two factors, $GSLM/Q^2Z_m$, and g/Z_g. A typical family of curves representing the first of these factors, for various values of the damping constant β, is plotted on logarithmic scales in Figure 2. Any curve in the family may be specified by the constants of its two asymptotes

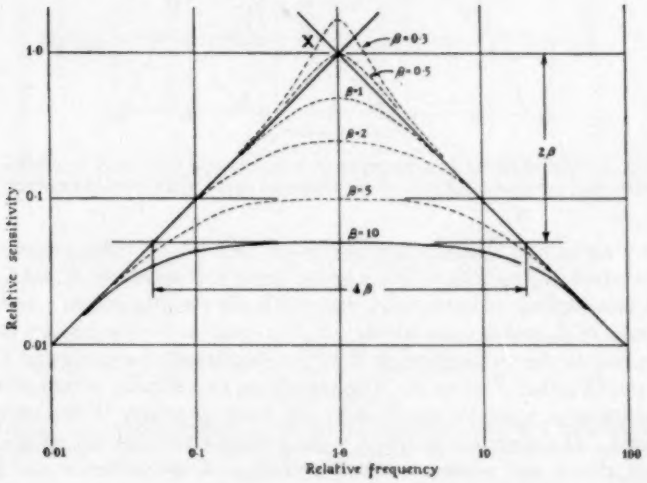


FIG. 2.—Normalized response curves for a simple harmonic oscillator with various degree of damping.

(which have slopes of $+1$ and -1 respectively) and the horizontal tangent. For values of β greater than 1, the three straight lines form quite a close approximation to the actual shape of the curve. If the asymptotes intersect at the point X on Figure 2 then the abscissa of X is the natural frequency of the instrument and the

level of the horizontal tangent corresponds to a sensitivity $1/2\beta$ times the ordinate of X. The length of the horizontal segment corresponds to a frequency ratio of $4\beta^2$.

The process of multiplication corresponds to the addition of two logarithms, so the sensitivity of a complete seismograph at any given frequency is obtained by adding together the ordinates which represent the galvanometer and seismometer sensitivities at that frequency. Within the limits of the trilinear approximation, we are adding together functions whose gradients are either ± 1 or zero, and the result, in the most general case, is a curve with five linear segments, having gradients ± 2, ± 1 or zero. A typical example of the general curve, with turning points at E, F, G and H is shown in Figure 3. The segments EF, FG and GH correspond, to the frequency bands over which the magnification, velocity sensitivity and acceleration sensitivity are approximately constant. The problem of designing a seismograph for a specific purpose therefore reduces to that of placing the turning points at the proper frequencies, and to the provision of sufficient electrical power to set the whole sensitivity curve at the desired level.

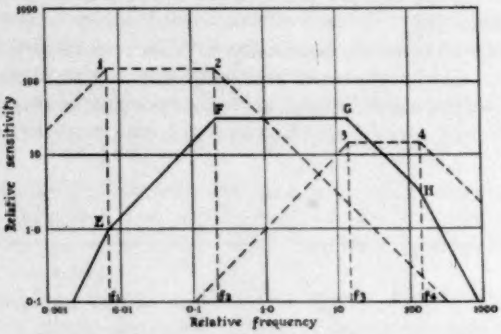


FIG. 3.—Asymptotes and tangents of the response curves of a seismometer and galvanometer, and of the response curve of the coupled system.

Figure 3 includes a trilinear approximation to a typical galvanometer characteristic, for which the turning points 1 and 2 occur at frequencies f_1 and f_2. It also includes a seismometer characteristic, for which the turning points 3 and 4 occur at frequencies of f_3 and f_4 respectively. If f_1 is taken to be the highest frequency (corresponding to the turning point E of the combined characteristic) f_2 can be chosen to match either F, G or H. The remaining two turning points of the combined characteristic must be matched by the turning points of the seismometer characteristic. The cases for which f_3 is the highest frequency repeat the form of those given above, and represent the interchange of seismometer and galvanometer constants. The six combinations correspond to the six values of ρ which satisfy equation (13).

Although all the solutions give identical response characteristics, the real seismographs which correspond to the solutions are far from being equally easy or economical to construct. In general, combinations which involve the use of seismometers or galvanometers of very long period are undesirable, partly because of instability of zero, and partly because, in the case of the galvanometer, the longest periods are only obtained by the use of large and heavy coils which require large

amounts of power to produce a given deflection. The exclusion of very long-period elements from the seismograph also tends to exclude very short-period elements, for the equations require that the natural frequencies of the seismometer and galvanometer must always fall on opposite sides of the sensitivity peak of the complete seismograph. We may therefore suspect that some of the most valuable combinations may be those in which the natural frequencies of the seismometer and galvanometer are close together near the centre of the combined characteristic, and in which the desired bandwidth is obtained by means of heavy damping.

We shall now investigate the design of the seismograph to have a constant velocity sensitivity over a very wide range of frequency. We shall set the turning points *F* and *G* to correspond to earth periods of 1 second and 400 seconds respectively. We could imagine such an instrument to be set up like a long-period Benioff combination, with a short-period seismometer and long-period galvanometer, both critically damped. In order to provide the desired periods for the turning points *F* and *G*, the seismometer and galvanometer periods would have to be 0·5 s and 800 s respectively, and the turning points *E* and *H* would fall at 0·25 s and 1 600 s respectively (see Figure 4).

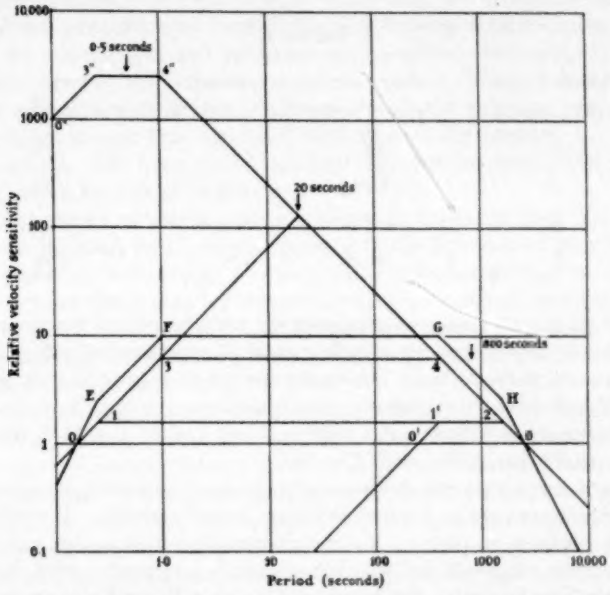


FIG. 4.—Asymptotes and tangents for the response curve of a wide-band seismograph, and for various galvanometers and seismometers which would produce the desired response.

The long–short combination would be extremely difficult to realize in practice. Alternative combinations are to set the turning points of the seismometer and galvanometer curves either at *E* and *G* and at *F* and *H* respectively, or to set them at *E* and *H* and at *F* and *G* respectively. The first alternative requires seismometer and galvanometer periods of 10 s and 40 s respectively, and $\alpha = \beta = 20$, whereas the second requires both periods to be set at 20 s, with damping constants of 10 and 40.

To check our approximate conclusions, we work out the constants A, B and C in accordance with equation (12). The results are given in Table 1, from which we see that whereas the 10–40 and 20–20 period combinations yield almost identical values of the seismograph constants, the values of A and C differ by about 20 per cent from those of the 0·5–800 combination. Algebraic analysis shows, in fact, that the exact match for the 0·5–800 seismograph is obtained when $\rho = 1$ and $\alpha = \beta = 20{\cdot}01$, and that the 20–20 and 10–40 combinations of Table 1 can be matched by a 0·5–800 combination for which $\alpha = \beta = 1{\cdot}25$. The discrepancy arises from the fact that the three segments of the trilinear approximation are not equally close to the exact characteristic, for the central segment is a tangent whereas the others are asymptotes. To embody this fact into a general rule, we note that if the response of a critically damped element is to be simulated by the interaction of an overdamped seismometer and galvanometer, then the turning points in the characteristics of the individual elements should coincide rather than being separated in frequency by two octaves. The most remarkable aspect of this conclusion is that such wide differences in the characteristics of the individual elements can so nearly compensate each other in the combined characteristic.

Table 1

Period combination	0·5–800	10–40	20–20
Damping constants	$\alpha = \beta = 1$	$\alpha = \beta = 20$	$\alpha = 40, \beta = 10$
ρ	40	2	1
A	80·05	100	100
B	1604	1604·25	1602
C	80·05	100	100

To show up the advantage of obtaining the wide frequency response by means of high damping rather than by a wide spread of seismometer and galvanometer period, we return to the trilinear curves for the 20–20 approximation as shown in Figure 4. The line O–O represents unit sensitivity, O–1–2–O is the response curve for a 20-second galvanometer with $\alpha = 40$, and O–3–4–O is the response of the 20-second seismometer with $\beta = 10$.

Let us now investigate the characteristics for an 800-second galvanometer and a 0·5-second seismometer to produce the same overall response. If we had chosen the thinnest available suspension for our 20-second galvanometer, the change to the characteristic required for the 800-second instrument would have to be obtained entirely by increasing the moment of inertia of the coil, and we would count ourselves fortunate if we could do this without having to use a stiffer suspension to carry the larger coil. Since the moment of inertia will increase as the square of the period, the 800-second coil will have 1 600 times the inertia of the 20-second one. The transducer constant g required for critical damping will be the same as that required to make $\alpha = 40$ in the 20-second instrument. As the sensitivity of a galvanometer at very long periods depends only on the transducer constant and the torsional constant of the suspension, the long-period asymptote to the sensitivity curve must remain along the line 2–O, and the complete characteristic for the 800-second galvanometer will be defined by the lines O′–1′–2–O.

If the seismometer characteristic is to add to the curve O′–1′–2–O to produce the curve O–E–F–G–H–O, the long-period asymptote of the seismometer characteristic must remain on the line 4–2–O. Extending this to the short-period end of the frequency range, we see that the seismometer characteristic must be the set of lines O″–3″–4″–O. Now, at very high frequencies, the outputs from a set of seismometers of similar design, but of different size, will be proportional to the constant G. As the line O″–3″ corresponds to 1 600 times the sensitivity of the line O–3, we find that the constant G for the 0·5-second seismometer must be 1 600 times greater than that of the 20-second seismometer. As the transducer constant depends on the square root of the mass of the energizing magnet, the 0·5–800-second combination will require $2{\cdot}56 \times 10^6$ times the magnetic material which is needed by a 20–20 seismograph of the same sensitivity. Even after one has allowed for the fact that some attenuation has to be introduced into the 20–20 circuit to keep the reaction down to an acceptable level, the 20–20 system remains overwhelmingly superior to the 0·5–800 combination.

Apart from its other advantages, the 20–20 system is very easy to set up and calibrate. The ease of setting up derives from the fact that both components settle into equilibrium comparatively rapidly when they are disconnected from each other, and this property also renders the step-by-step determination of the constants of the seismometer and galvanometer comparatively rapid. If the calibration is carried out by the sine-wave simulator method (Murphy & others 1954) or by means of a Maxwell Bridge (Willmore 1959), the inherent symmetry of the response curve is such that one need only observe the response for periods less than 20 seconds. The long-period response may then be determined by reflecting the curve across the line of symmetry.

The advantages of placing the seismometer resonance near the centre of a desired pass-band are by no means confined to the detection of very long periods. Thus in planetary seismology, or for work in inaccessible regions of the Earth, requirements are developing for compact and robust instruments to cover ordinary earthquake frequencies. A $\frac{1}{2}$-second system with $\alpha = \beta = 20$ might prove to be extremely suitable for installation in a difficult region, and would cover the range of periods from 1/80th s to 20 s.

As a second illustration of the method, we shall consider the design of a seismograph with constant magnification over the range of periods from 1 to 20 s, and a constant velocity sensitivity from 20 to 200 s period. The required characteristic is shown in Figure 5. One solution is evidently to use a 400-second galvanometer, critically damped, to obtain the long-period roll-off, and to drive this galvanometer with a seismometer having a period of 4·5 s and a damping factor of about 2·25 times critical. An alternative method is to set the turning points of the seismometer characteristic at 1 s and 400 s respectively, and to set the galvanometer turning points at 20 s and 400 s. The latter combination requires a 20-second seismometer with $\beta = 10$, and a 90-second galvanometer with $\alpha = 2{\cdot}5$. Once again, we see that the combination which has the higher damping coefficients avoids the use of the very long-period galvanometer. The figure includes the response curves computed algebraically for $\lambda = 1$ and $\lambda = 0{\cdot}75$.

To meet the requirements for long-period surface-wave studies, some workers are suggesting that the magnification should be held to a constant or rising level out to periods considerably longer than 20 s. However, we see from Figure 5 that the 20-second turning point (which is derived from the galvanometer response curve) cannot be pushed out to a longer period unless the long-period turning point

of the galvanometer response is brought into a shorter period. The best that can be done to broaden the range of constant magnification is to increase the seismometer damping constant to a very large value. This will push the seismometer turning points out to very high and very low frequencies, and the overall response will approximate to that of a single pendulum having the same period and damping constant as the galvanometer.

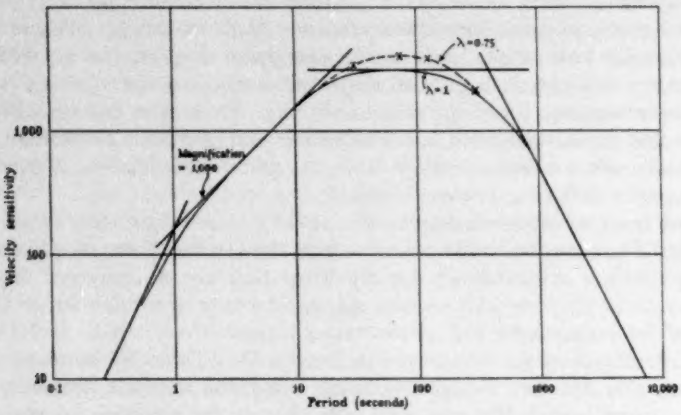


FIG. 5.—Response curve of a seismograph designed to yield approximately constant magnification for earth periods between 1 second and 20 seconds. The solid lines show the response when the reaction of the galvanometer is negligible, and the dotted curve shows the response when $\lambda = 0{\cdot}75$.

4. Effect of feedback through an amplifier

In the foregoing discussion, we have surveyed the theory of passive seismographs, which are operated entirely by the energy which can be extracted from the ground. In recent years, however, there has been a trend towards the use of electronic amplifiers in seismometry, and when such amplifiers are introduced into the circuit we have the opportunity of feeding back the amplifier output into the seismograph circuit and thereby modifying the response.

At first it might be thought that a very large number of possibilities would have to be considered, for the feedback circuits can be set up in many different ways, and each circuit can be set up in conjunction with any of the possible combinations of seismometer and galvanometer.

In fact, the possibilities are somewhat limited. Any feedback current which is injected into the seismograph circuit will divide itself between the seismometer and galvanometer coils in a certain ratio, and we can consider the effects on the two elements of the system separately. We shall assume that the feedback current will be derived from a device such as a photoelectric amplifier whose output is proportional to the deflection of the galvanometer spot, and we shall consider the effects of resistance-capacity coupling networks which will feed back the first or second differential of the amplifier output. It seems likely that most useful cases will be covered by a discussion on these lines, although the extension of the argument to cover other types of feedback would not be very difficult.

First, let us consider the effect of that part of the feedback current which flows in the galvanometer circuit. Let the feedback current be denoted by the quantity $p\theta$. Then the equation of motion of the galvanometer becomes

$$k\frac{d^2\theta}{dt^2}+d\frac{d\theta}{dt}+u\theta = g(i+p\theta). \tag{16}$$

Evidently, the quantity $gp\theta$ can be transferred to the left-hand side of equation (16), where it will be subtracted from the torsional couple $u\theta$. If the feedback network had incorporated differentiating elements to yield feedback proportional to $d\theta/dt$ or $d^2\theta/dt^2$ the transfer to the left-hand side of equation (16) would modify the effective damping constant or the effective moment of inertia of the coil. Thus, although these types of feedback may provide a convenient method of modifying the characteristics of the galvanometer, they do not produce types of response which are not obtainable, at least theoretically, in passive systems.

On passing to the effect of feedback into the seismometer, we shall first consider the case in which the amplifier output is differentiated once, giving a feedback current of $P(d\theta/dt)$. The equation of motion of the seismometer is

$$K\frac{d^2\phi}{dt^2}+D\frac{d\phi}{dt}+u\phi = LM\frac{d^2x}{dt^2}-G\left(I+P\frac{d\theta}{dt}\right).$$

Making the substitutions which previously led to the reduced equation of motion (equation (5)) we find

$$\frac{d^2\phi}{dt^2}+2\beta\Omega_0\frac{d\phi}{dt}+\Omega_0{}^2\phi = \frac{1}{l}\frac{d^2x}{dt^2}+\left(2\beta\Omega_0\sigma_s-\frac{PG}{K}\right)\frac{d\theta}{dt}, \tag{17}$$

$$= \frac{1}{l}\frac{d^2x}{dt^2}+2\beta\Omega_0\sigma_s'\frac{d\theta}{dt}. \tag{18}$$

where

$$\sigma_s' = \sigma_s-\frac{PG}{2K\beta\Omega_0}.$$

We can now follow through the entire computation of seismograph response using σ_s' in place of σ_s. The result will be the same as before, except that the constant λ which expresses the galvanometer reaction is no longer subject to the restriction $1 > \lambda > 0$.

The possibility of making $\lambda > 1$ leads to an interesting extension of the range of characteristics obtainable. We illustrate this point by returning to the equation of the wide-band seismograph, which was made up from a seismometer and galvanometer having equal periods and a high degree of damping. The response of this combination may be derived from equation (10) by setting $\omega_0 = \Omega_0 = \omega'$. Let us now focus our attention in the high-frequency end of the response curve by setting $\omega \gg \omega'$, so that the equation approximates to

$$\frac{\Theta}{(dX/dt)} = \frac{-2\alpha\sigma_g}{l\omega\{(\omega/\omega')-2j(\alpha+\beta)-4\lambda\alpha\beta(\omega'/\omega)\}}. \tag{19}$$

Equation (19) has the same form as the response of a long-period galvanometer driven by a seismometer of short period, where the natural frequency of the seismometer is given by $(4\lambda\alpha\beta)^{\frac{1}{2}}\omega'/2\pi$, and the damping constant is $(\alpha+\beta)/(4\lambda\alpha\beta)^{\frac{1}{2}}$.

If we had considered the low-frequency end of the response curve, we would have seen that it has the same form as that due to a galvanometer with a natural frequency of $\omega'/2\pi(4\lambda\alpha\beta)^{\frac{1}{2}}$ and a damping constant of $(\alpha+\beta)/(4\lambda\alpha\beta)^{\frac{1}{2}}$.

In passive systems, for which λ is always less than unity, the ends of the response curve can never simulate the behaviour of components having less than critical damping. When the possibility of feedback through an amplifier is introduced, the restriction on λ is removed, and the behaviour of underdamped components can be imitated. As the three constants α, β and λ enter into the overall response only through the combinations $(\alpha+\beta)$ and $\lambda\alpha\beta$, one of the three variables can be fixed arbitrarily within wide limits, and any desired response can be obtained by adjustments to the other two.

There is another way of considering the effect of feedback into the seismometer coil, which depends on the fact that the effect of a ground acceleration equal to d^2X/dt^2 is equivalent to the application of a force to the seismometer boom, such that the moment of the force about the hinge is $LM(d^2X/dt^2)$. The feedback current exerts such a force, so the current simulates the effect of a proportionate acceleration of the seismometer pier.

If the feedback current, which is proportional to $d\theta/dt$, simulates an acceleration of the ground, it follows that the galvanometer deflection θ simulates velocity. Thus we can regard our entire seismograph as a "black box" with earth velocity as the input, with an output proportional to θ, and with velocity proportional to θ fed back to the input. The flattening of the velocity sensitivity curve is now recognized as the natural consequence of the application of negative velocity feedback. The peaks which may appear at the high-frequency and low-frequency ends of the pass-band are due to the phase shift which occurs at these frequencies and are therefore closely analogous to the peaks which are a familiar feature of audio-frequency feedback amplifiers. We see also that the galvanometer reaction in passive systems is equivalent to positive feedback of the earth velocity, and that the tendency of the reaction to produce a central peak in the velocity sensitivity characteristic is a natural consequence of such feedback.

Let us now consider the effect of feeding back a current equal to $P\theta$, or to $P(d^2\theta/dt^2)$. On inserting the appropriate term in place of $P(d\theta/dt)$ in equation (18) we solve the equations of motion as before. When the feedback is proportional to θ, the solution is

$$\Theta\left(\frac{dX}{dt}\right)^{-1} = -2\alpha\sigma_g(l\Omega_0)^{-1}\times$$
$$\times\left\{\frac{\omega^2}{\omega_0\Omega_0}+\frac{\Omega_0\omega_0}{\omega^2}-\frac{\omega_0}{\Omega_0}-\frac{\Omega_0}{\omega_0}-4\lambda\alpha\beta-\right.$$
$$\left.-j\left[2\alpha\left(\frac{\omega}{\Omega_0}-\frac{\Omega_0}{\omega}\right)+2\beta\left(\frac{\omega}{\omega_0}-\frac{\omega_0}{\omega}\right)-\frac{PG}{\omega\omega_0\Omega_0}\right]\right\}. \qquad (20)$$

Here we see that the feedback introduces a term in $1/j\omega$ into the denominator of the expression for the sensitivity. The presence of this term will extend the range of frequencies over which the term in $1/j\omega$ controls the denominator, which is the range over which the seismometer yields approximately constant response to

an acceleration of the ground. Following the argument which was applied when the feedback was proportional to $d\theta/dt$, we recognize the flattened acceleration response as a natural consequence of negative acceleration feedback.

When the feedback is proportional to $d^2\theta/dt^2$, the solution is

$$\Theta\left(\frac{dX}{dt}\right)^{-1} - 2\alpha\sigma_g\,(l\Omega_0)^{-1}\times$$
$$\times\left\{\frac{\omega^2}{\omega_0\Omega_0}+\frac{\Omega_0\omega_0}{\omega^2}-\frac{\omega_0}{\Omega_0}-\frac{\Omega_0}{\omega_0}-4\lambda\alpha\beta-\right.$$
$$\left.-j\left[2\alpha\left(\frac{\omega}{\Omega_0}-\frac{\Omega_0}{\omega}\right)+2\beta\left(\frac{\omega}{\omega_0}-\frac{\omega_0}{\omega}\right)+\frac{PG\omega}{\omega_0\Omega_0}\right]\right\}. \quad (21)$$

In this case, the feedback term broadens the range of constant magnification, as we should expect from the application of displacement feedback.

5. Performance of an experimental prototype

The foregoing conclusions may be illustrated by considering the performance of a seismograph which has been set up at the Dominion Observatory in Ottawa. This instrument consists of a long-period Sprengnether seismometer with two transducer systems, each containing a coil wound to a resistance of 100 000 ohms. The seismometer period is adjusted to 20 s and the power of the transducers is such that the pendulum damping is about four times critical when the coils are short circuited.

One of the seismometer coils is connected to a Tinsley type 4789 galvanometer, having a coil resistance of about 1 000 ohms and a critical damping resistance of about 60 000 ohms. A shunt of 5 000 ohms across the galvanometer terminals makes $\alpha \simeq 10$.

The galvanometer is operated with an optical beam of 50 cm. The optical system is designed to produce a rectangular light spot, which is split between a pair of vacuum photocells. The photocells feed into the grids of a 12 AU 7 tube wired as a pair of cathode followers, which is followed by a second 12 AU 7, used as a power amplifier. The output of the second tube is about 7 volts per millimetre of the light-spot deflection, and is approximately linear for deflections of the light-spot up to about 15 mm on each side of zero. Feedback is taken through a 0·03 μF mica condenser into the second coil of the seismometer and is set at an appropriate value by means of a shunt resistor of 10 000 ohms (see Figure 6).

The results of a calibration experiment which was carried out over the short-period end of the frequency range are plotted against an arbitrary velocity sensitivity scale on Figure 7. The points may be compared with the theoretical curves for 20–20 seismographs with $\alpha = 10$, $\beta = 4$ and $\lambda = 10$, and with $\alpha = \beta = 20$, $\lambda = 1$. The constants used in drawing the first of the two theoretical curves are close to those which have been used in the actual seismograph, and it can be seen that a slight adjustment of the values would produce an almost perfect fit between the observed and calculated results. The output level which corresponds to the flat central part of the calibration curve is 30 000 V/cm/s, which means that the system has ample power for working into visual or electronic recording systems.

The prototype equipment was assembled from parts which are readily obtainable commercially, many of which were ordered before the principles of the design

were fully understood. It is now clear that substantial improvements in performance could be achieved by the use of more effectively matched components.

One of the most important changes would be to use a galvanometer with a stronger magnet, because the Tinsley 4789, in common with most long-period instruments, is designed to operate with quite a weak field in the gap. By using a magnet with enough power to saturate the pole pieces, the sensitivity would be increased by a factor of 7, and the damping constant α would increase by a factor of 50. A further improvement could be obtained by eliminating the mis-match between the 100 000-ohm seismometer coil and the 1 000-ohm galvanometer suspension. A system re-designed on these lines would be somewhat more susceptible to disturbance by thermal e.m.f.s than the present one, but these could be reduced

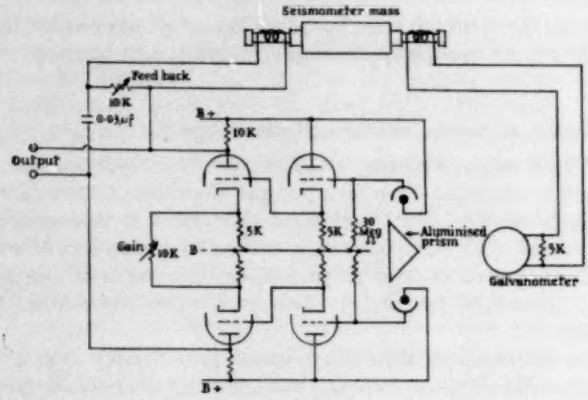


FIG. 6.—Photo-electric amplifier and feed-back circuit for a wide-band velocity seismograph.

by more attention to the details of the design. It seems certain that the combined effects of design improvements could add at least one decade of frequency response to each end of the pass-band, without involving any loss in signal–noise ratio within the detecting system. However, the performance of the equipment already appears to be limited by the effects of thermal fluctuations in the vault and by instability of the pier, and a much more stable environment will be required before we can expect to utilize increased sensitivity outside the limits of the present pass-band.

6. Use of detectors other than galvanometers

The preceding argument has been built on the assumption that the seismometer current would be detected by a moving-coil galvanometer. In fact several other types of detectors are available, and some have been used by other workers to obtain wide-band response characteristics (see, for example, Benioff 1960). We should therefore ask ourselves how the advantages and disadvantages of galvanometers compare with other available detectors.

The galvanometer characteristic as portrayed in Figure 2 is that of a current integrator, over the frequency range for which the mechanical impedance of the

suspension is dominated by the damping term. At frequencies remote from the centre of the pass-band the response falls below that of an ideal integrator at the rate of about 6 db per octave, and phase-shifts which approach the values of $\pm\pi/2$ at very high and very low frequencies respectively, are introduced. It is these phase shifts, added to those introduced by the seismometer, which generate the peaks at the ends of the pass-band when large amounts of negative velocity feedback are introduced, but the fact that the total phase shift round the feedback path can never quite reach the critical value of π radians means that the system can never break into self-sustained oscillation. The desirable phase-shift characteristic of a moving-coil galvanometer is combined with an ability to detect e.m.f.s of the order of 10^{-8} V in circuits of moderate impedance.

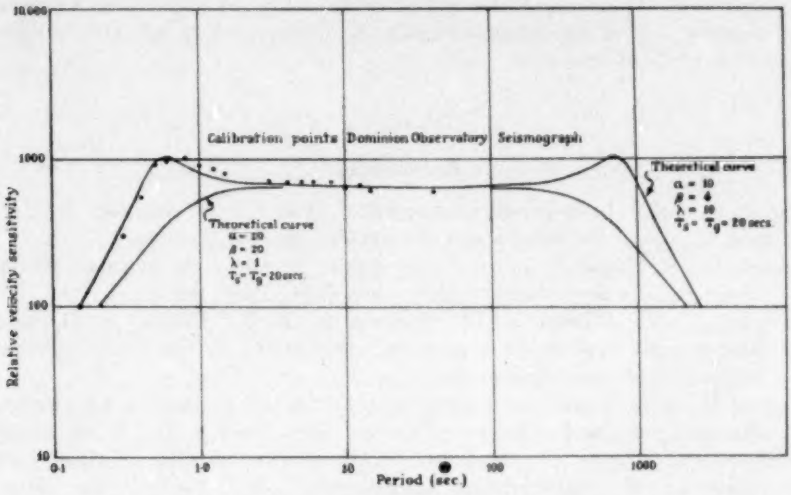


FIG. 7.—Theoretical curves and experimental calibration points for a wide-band seismograph.

Against these advantages, we must set the disadvantage that a sensitive galvanometer amplifier is a bulky and somewhat fragile object. For extreme low-frequency use, it might be thought that chopper-stabilized or magnetic amplifiers would provide effective substitutes for galvanometers, but the difficulty with many of such instruments is that they contain carrier-rejection filters which can introduce sufficient phase-shift to cause oscillation when the feedback loop is closed. This difficulty might be avoided by the use of a carrier frequency high enough to move the phase-shifting property of the rejection filter out of the frequency band within which mechanical oscillations of the seismograph are liable to occur.

Another point to consider when one is aiming for a flat response to earth velocity is that the integrating property of the over-damped galvanometer is not an essential feature of the feedback loop; on the contrary, it is a property which has to be eliminated by the use of a differentiating circuit. Instead, one could use a flat amplifier with controlled roll-off at the ends of the desired pass-band, and apply the output of this amplifier to the seismometer through a d.c. feedback network. The amplifier output would then be proportional to the *acceleration* of

the ground over a wide frequency range, and flat velocity response could be obtained by the use of a single-stage integrator outside the feedback loop.

A galvanometer is a less desirable circuit element when the ultimate objective is to obtain a flat response to displacement or acceleration of the ground instead of to the velocity. The reason is that the response of a seismometer–galvanometer combination to either of these parameters involves a phase shift which approaches $3\pi/2$ radians at one end of the pass-band. A phase shift of π radians will therefore exist at some frequency within the pass-band, and the application of more than a small amount of feedback may cause the system to oscillate.

To sum up, it would appear that the moving-coil galvanometer is a very valuable element in feedback seismographs for use in fixed stations, particularly when a wide-band response to ground velocity is required. Other types of detectors may be used when robustness and portability are prime requirements, or when the objective is to obtain constant response to displacement or acceleration over a very wide range of frequency.

References

Benioff, H., 1960. Long-period seismographs. *Bull. Seismol. Soc. Am.*, **50**, 1.

Galitzin, B., 1914. *Vorlesungen über Seismometrie Teubner* (Leipzig).

Grenet, G., & Coulomb, J., 1935. Nouveaux principes de construction des seismographes électromagnetiques. *Ann. Phys.*, Paris, Sér. 11, **3**, 321.

Murphy, L. M., Wilson, R. M., Burgess, L. R. K., Pearce, T. H. 1954. Response curves of an electromagnetic seismograph by sine-wave simulator method. *Bull. Seismol. Soc. Amer.*, **44**, 7.

Wenner, F., 1929. A new seismometer equipped for electromagnetic damping and electromagnetic and optical magnification. *Bur. Stand. J. Res.*, Wash., **2**, 963.

Willmore, P. L., 1959. The application of the Maxwell impedance bridge to the calibration of electromagnetic seismographs. *Bull. Seismol. Soc. Amer.*, **49**, 99.

the ground over a wide frequency range, and flat velocity response could be obtained by the use of a single-stage integrator outside the feedback loop.

A galvanometer is a less desirable circuit element when the ultimate objective is to obtain a flat response to displacement or acceleration of the ground instead of to the velocity. The reason is that the response of a seismometer–galvanometer combination to either of these parameters involves a phase shift which approaches $3\pi/2$ radians at one end of the pass-band. A phase shift of π radians will therefore exist at some frequency within the pass-band, and the application of more than a small amount of feedback may cause the system to oscillate.

To sum up, it would appear that the moving-coil galvanometer is a very valuable element in feedback seismographs for use in fixed stations, particularly when a wide-band response to ground velocity is required. Other types of detectors may be used when robustness and portability are prime requirements, or when the objective is to obtain constant response to displacement or acceleration over a very wide range of frequency.

References

Benioff, H., 1960. Long-period seismographs. *Bull. Seismol. Soc. Am.*, **50,** 1.

Galitzin, B., 1914. *Vorlesungen über Seismometrie Teubner* (Leipzig).

Grenet, G., & Coulomb, J., 1935. Nouveaux principes de construction des seismographes électromagnetiques. *Ann. Phys.*, Paris, Sér. 11, **3,** 321.

Murphy, L. M., Wilson, R. M., Burgess, L. R. K., Pearce, T. H. 1954. Response curves of an electromagnetic seismograph by sine-wave simulator method. *Bull. Seismol. Soc. Amer.*, **44,** 7.

Wenner, F., 1929. A new seismometer equipped for electromagnetic damping and electromagnetic and optical magnification. *Bur. Stand. J. Res.*, Wash., **2,** 963.

Willmore, P. L., 1959. The application of the Maxwell impedance bridge to the calibration of electromagnetic seismographs. *Bull. Seismol. Soc. Amer.*, **49,** 99.

Geophysical Journal

RECOMMENDATIONS TO AUTHORS

1. Authors are advised to follow the general suggestions for the preparation of scientific papers given in *Notes on the preparation of papers to be communicated to the Royal Society* and in *The printing of mathematics* by T. W. Chaundy, P. R. Barrett and C. Batey (Oxford University Press, 1954). In particular the abbreviations for the names of periodicals used in the *Geophysical Journal* are those given in *Notes on the preparation of papers* . . . and the symbols, signs and abbreviations are those recommended for British scientific publications by the Symbols Committee of the Royal Society.

2. Papers should be typewritten with double spacing and on one side of the paper only. Two copies, one of them the top copy, should be submitted. All papers must be accompanied by a summary; summaries are not required for *Letters to the Editors*.

3. Tables are printed without rules. They should be numbered serially with Arabic numerals. Table headings should be brief. Units should be placed at the head of the column. Tables should be typed on separate sheets and their positions in the text indicated on the copy.

4. Illustrations, especially photographs, should be kept to a minimum, and the same information should not in general be given in both tables and illustrations. Line drawings should be in dense black ink on smooth white board or transparent tracing film. Original drawings must not be lettered; lettering should be indicated on copies or photoprints which must be supplied in addition to the originals. Wherever possible, lettering should be kept outside the diagram so that it may be set in type. Diagrams should be drawn at twice to three times the size at which they will be printed. The maximum dimensions for diagrams, when reduced and with their lettering and captions, are 8 in by 5 in. A wide variety of flat tints of dots, lines and shadings can be applied to line drawings by the blockmaker. Photographs for reproduction should be unmounted glossy prints and should be accompanied by lettered prints.

5. References should be quoted in the form recommended by the Royal Society but authors may also quote titles of papers if they wish. References in the text are made in the form (Smith 1940).

6. Authors whose work involves much mathematics should follow carefully the recommendations in *The Printing of Mathematics*.

7. The Editors will be glad to advise authors on any special points of difficulty arising in the preparation of papers for the *Journal*.